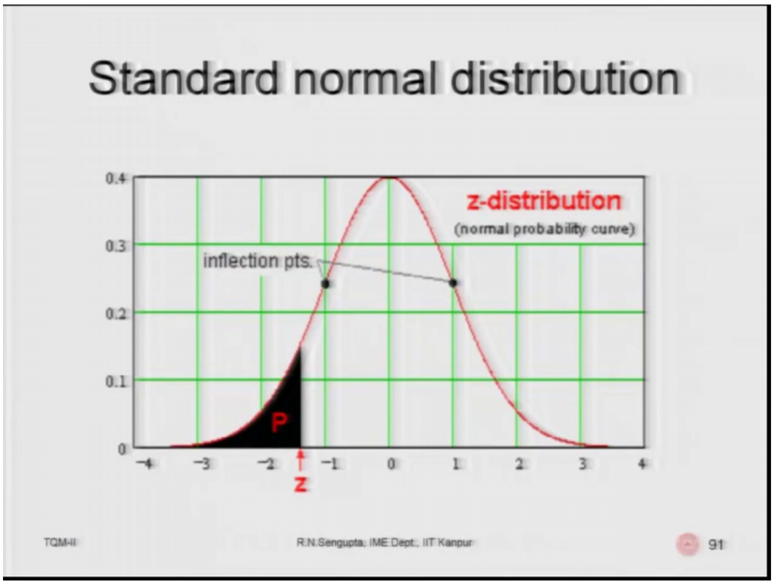
Total Quality Management-II
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 04
Distribution of a Random Variable-II

Welcome back my dear friends a very good morning good afternoon good evening to all the students and all the people who are taking this course. I am Raghunandan Sengupta from the IME department IIT Kanpur, and this is the fourth lecture for the TQM 2 and as you know that it will be more of a building block for that design of experiments which will cover within another 2 to 3 lectures.

So, we were discussing about normal distribution the relevance of normal distribution will come out later, we did mention that how if x is normally distributed with mean mu and sigma square as the variance, how it can be transformed into a z distribution with which is the standard normal and how the tables of standard normal can be utilized to do the calculations. And we also initially before that we understood the concept of p d f, c d f and that c d f would be utilized time and again in the calculations we will see to that in due course of time.

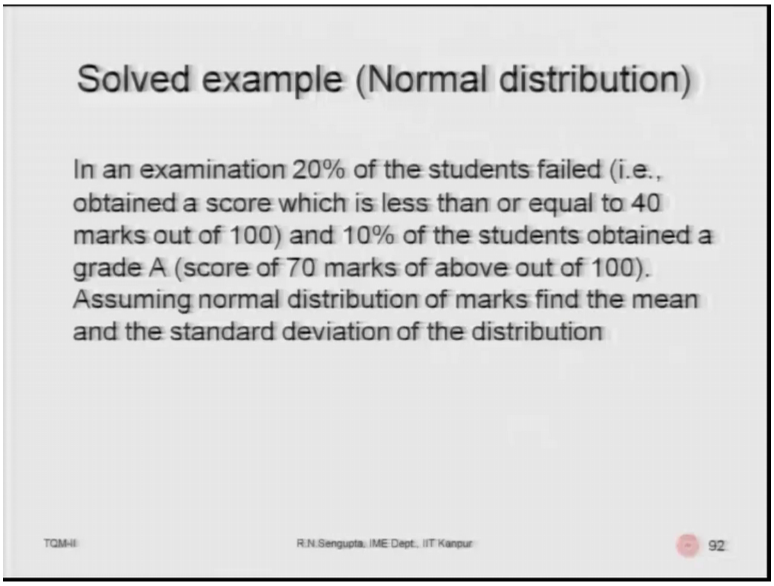
(Refer Slide Time: 01:20)



So, the standard normal distribution basically if you consider the mean value mode value or the median is 0, and as per the transformation and the standard deviation or the variance is 1. So, if you go if you are looking the distribution from your side if you go plus 1 sigma on to the right plus and minus 1 sigma on to the left, the overall coverage of the area you can find it from the standard normal table and the overall area covered would be about I am giving it approximate values it will be about 67 percentage.

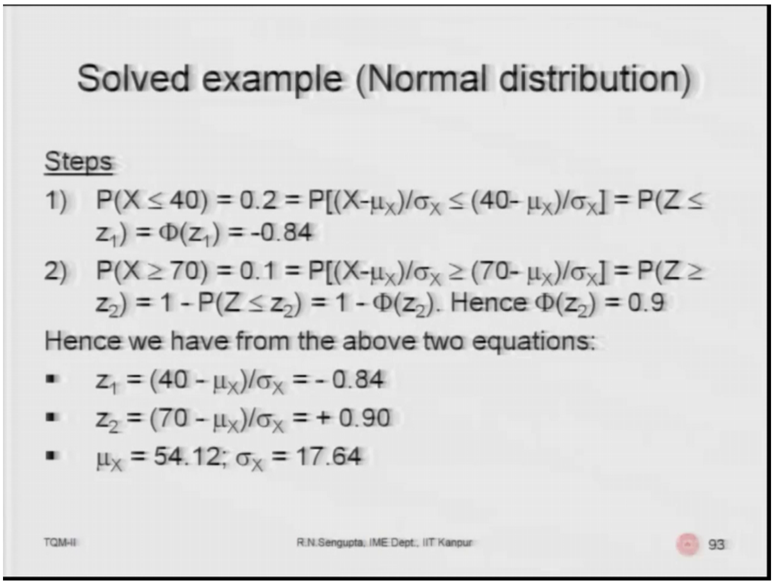
And those points where if you go sigma plus and sigma minus; that means, from the mean value those are the inflection points which means that the rate of change of the derivative basically takes place from negative to positive, positive or negative and we know point of inflections are the points where the second derivative is also 0. So, in maximization and minimization we know that the first derivatives are 0 and the second derivatives are respectively positive and so and so forth, based on which we can find out these values.

(Refer Slide Time: 02:00)



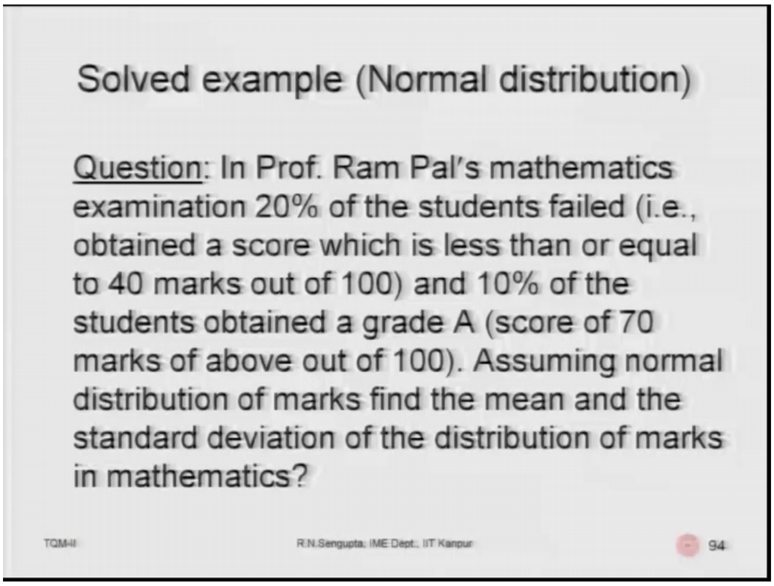
So, considering in a examination 20 percent of the students failed scored 40 percent marks less 10 percent of student obtained a grade 70 and more and. So, there are two sets of information from this process of information we need to find out the mean and the variance which means.

(Refer Slide Time: 02:43)



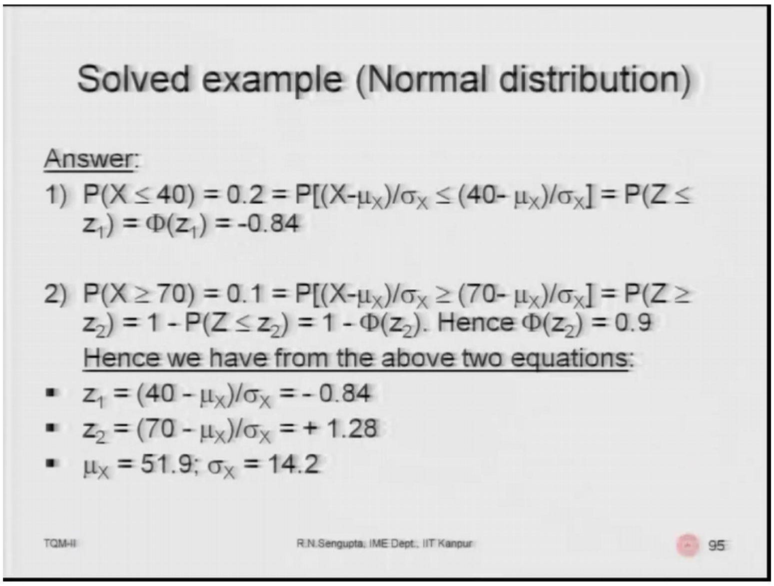
Two unknowns, two equations and we can solve them as shown they are very simple concept they would not be utilized that much in TQM 2, but I want you to go a little bit fast. So, you will understand the significance of standard normal table from there we can find out z 1 and z 2 and mu and solve the problems accordingly.

(Refer Slide Time: 03:01)

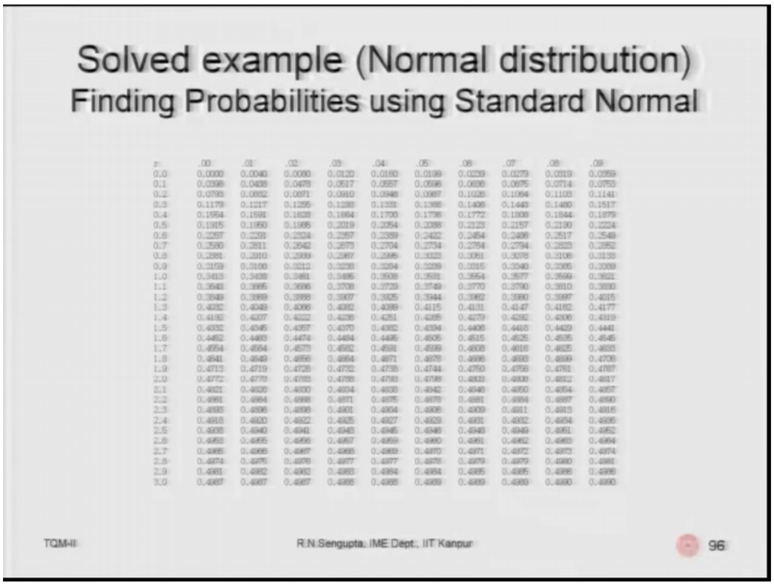


So, this is the same problem in professor ram pals mathematics examination 20 percent of the students failed obtained 40 percent marks and less 10 percent obtained a grade 70 percent marks and more you can find out the mean and the standard deviation; which comes out to be 51.9 and 14.2.

(Refer Slide Time: 03:14)

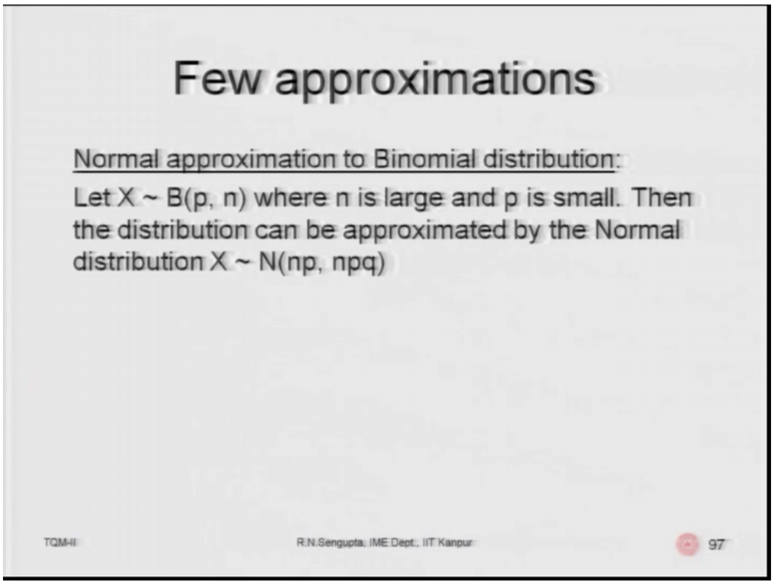


(Refer Slide Time: 03:21)



So, finding of the probabilities as I said the, of I am just repeating the first column consists of z values the topmost row consists of the decimal values of z and the in between values inside that the matrix whole matrix is basically the c d f values.

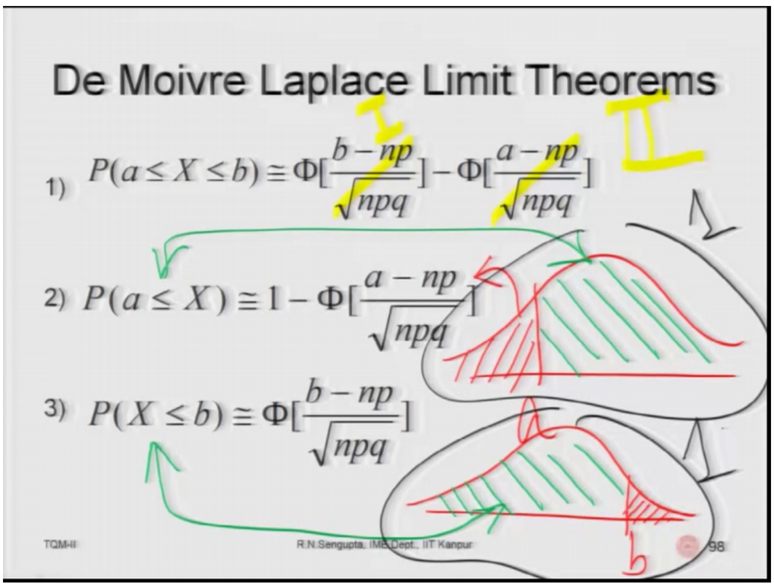
(Refer Slide Time: 03:40)



So, normal approximation to binomial distribution, so consider x is binomial distributed with parameters p n n p is the probability of success q is the probability of the failure such that q is equal to 1 minus p. So, if n f where f n is large and p is small and then the distribution can be approximated by the normal distribution with the mean value is as n p and the variance as n p q and you can convert that also into a standard normal table and do all your calculations.

So, this is what is given here, but De Moivre's Laplace limit theorem.

(Refer Slide Time: 04:13)



If you want to find out the probability of x being a random variable between a and b and m and x is basically a binomial distributed where n where n is very large p is very small and you convert to a normal distribution, then the values of the c d f values for till b and till a are given by the standard normal distribution. So, if how we convert the standard normal distribution is basically x minus mu divided by the standard deviation. So, here the standard deviation is basically square root of n p q and the mean value is n p and from that you can find out the values as given are like this.

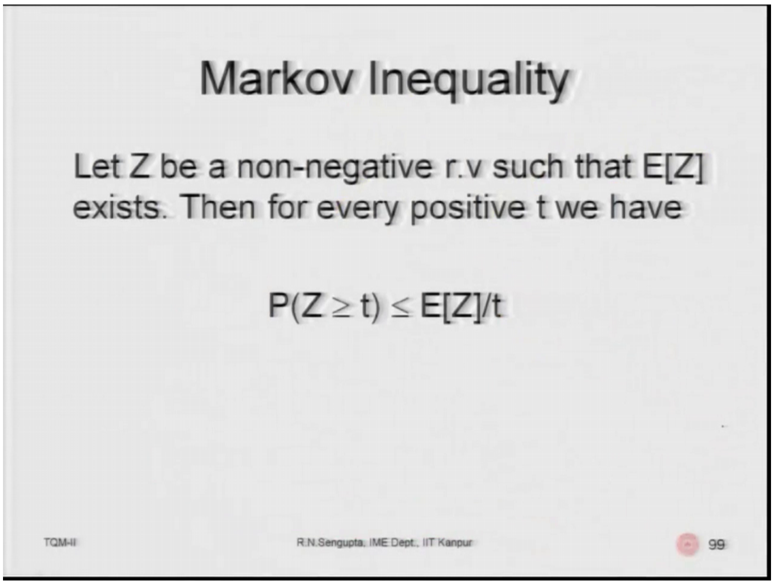
So, we have just highlighted, so this is the first one is basically a c d f value starting from minus infinity to be second values all the c d f values some of the c d f values from minus infinity to a and difference between that will be given the total area covered from a to b. Now if a b tends to positive infinity it means it basically the whole area becomes 1 so; obviously, the capital phi value would be 1.

So, you have to find out the probability from a of x being greater than a. So, technically the value would be 1 minus the overall sum of the c d f values from p d f values c d f values is the some of that from minus infinity to a. So, 1 minus that will give you the whole area. So, let me again I am sure everybody understanding it, but I am just highlighted. So, this is a, so this is this value. So, 1 minus 1 is the whole area. So, 1 minus this will be give me this area, so let me highlight it with another colour, so this is the value which I wanted.

Now, if I want to find out as a tends to infinity I mean minus infinity, so probability of x less than b. So, again the same distribution and draw it this is b. So, the area being given here, so I need to find out the area here this is basically this. So, you add up all from minus infinity to the value as given will be.

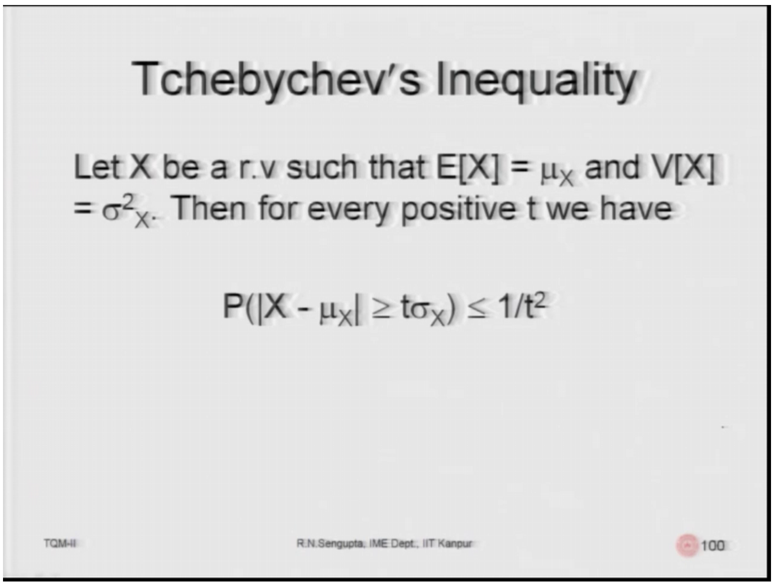
So, you can find it at accordingly 1, but; obviously, remember the whole area another distribution the whole area covered here is 1 whole area covered is 1; hence it can subtract from 1 or find out that integration of all these values.

(Refer Slide Time: 07:18)



So, Markov's inequality, so let Z be a non-negative random variables as that expected value of Z exists. So, it does not say anything about the variance is does not say they have anything about type of distribution, then for every positive t we will have basically the probability of Z being greater than that fixed value of t is less than equal to that is the bound is less than equal to the expected value of Z divided by t and this inequalities are utilized for different cases.

(Refer Slide Time: 07:49)

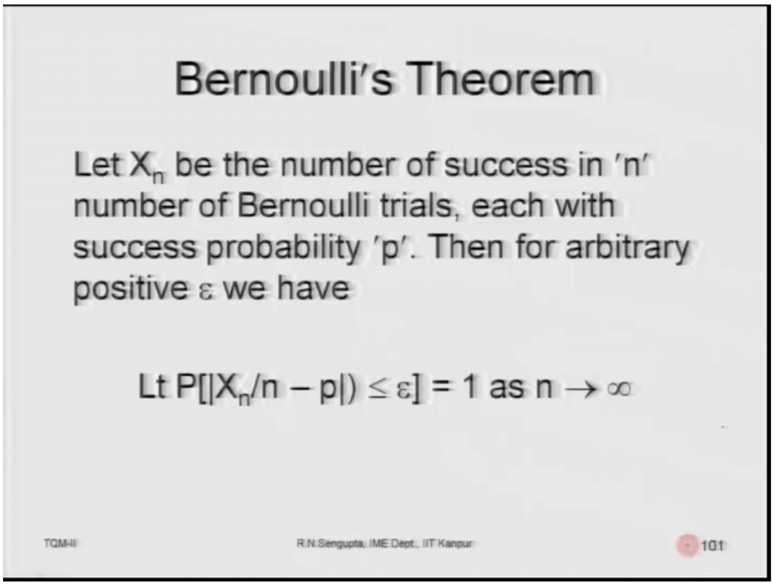


Now, let X be a random variable which is the Tchebychev's inequality let X be a random variables such that the expected value which is the first moment and the variance which is the same second moment exist, and then for every positive t and it doesn't mention anything about whether it is a positive random variable of or a positive a negative random variable.

Then the bound of the difference between x and its mean value would be greater than equal to some constant value into the standard deviation which should always be less than equal to the 1 by square of that that is constant value. So, consider very simply that if t is 1 which is 1 plus minus sigma. So, that would always be less than equal to 1 and you can find out the corresponding probabilities and the bound.

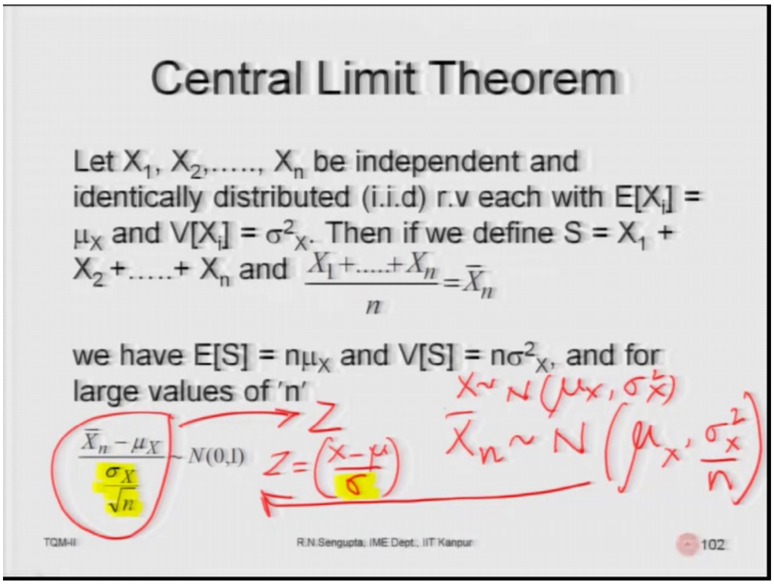
So, accordingly we will use them very rarely, but just as precursor then Bernoulli theorem. So, let X n be the number of successes in n number of Bernoulli trials each with the success probability p so; obviously, the unsuccessful 1 would be q.

(Refer Slide Time: 08:52)



Then for any arbitrary positive epsilon we will have limit of n tending to infinity the difference between the relative frequency of the probability would be as close as possible to 0 as n increases this is the basic essence. So, epsilon is basically a function of n epsilon n becomes epsilon n or epsilon whatever you see becomes smaller and smaller as n increases. So, which means the difference between the relative frequency of the chance and the probability basically tends towards 0, so this is what it actually means.

(Refer Slide Time: 09:26)



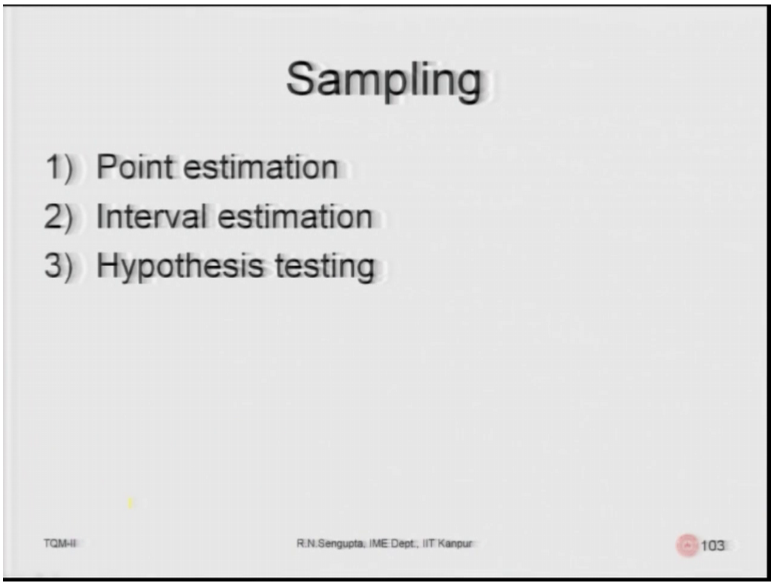
Now, the central limit theorem, so X let X 1 to X n be the independent and identically distributed i i ds random variables random variables each with the with expected value of X is suffix is being equal to mu suffix x, and the variances for each of them being sigma squares suffix x then if we define s as the sum.

So, sum means that is basically sum of X 1 X 2 X 3 till X n, they and we defined this average of this sum as x bar n, then we know that the expected value of s which is the sum is equal to n into mu x because each expected value they are independent i i ds. So, each expected value is mu x, so how many are such mu x s out there are n.

So, if you add them up it becomes n mu x and if I want to find out the variances of that it will be basically trying to find out. So, as their i i ds; obviously, the concept of co variances would not come. So, you will basically have n number of variances. So, n number of variances would be which you would find out the variances accordingly. So, these are given and you can find it out as n sigma square and for large n as becomes large n. So, x n x bar n and basically would have what I mean is actually this. So, let me use a coloured 1 this is normal with mu x sigma square x by n.

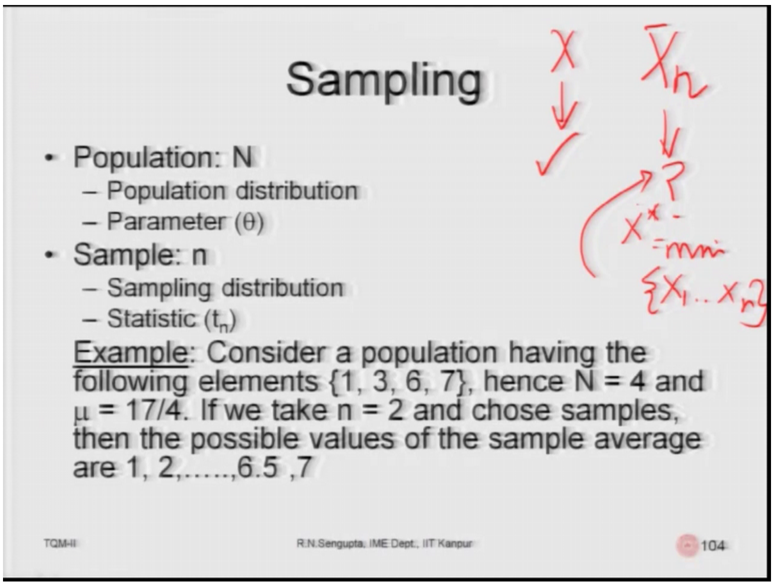
So, when you convert this you utilize this and you get the concept of the standard normal. So, this basically this is the standard normal, and in case when say for example, x is distributed for normal mu x sigma square x then in that case z is standard normal with this. So, the only differences which is happening which I should highlight is this 1 and this 1 in 1 case it is divided by square root of n and another case not because you will understand the calculations would give you accordingly.

(Refer Slide Time: 11:53)



So, we will consider the sampling, so there is point estimation interval estimation hypothesis testing. So, we will go slowly through point estimation interval estimation hypothesis testing.

(Refer Slide Time: 12:05)



So, considered is sampling you pick up a population you have actual of population of capital N, the population distribution is known you have a parameter theta, theta can be scalar or vector depending on the problem, and you have be basically pick up a sample, sample is a size n and the sample distribution is given.

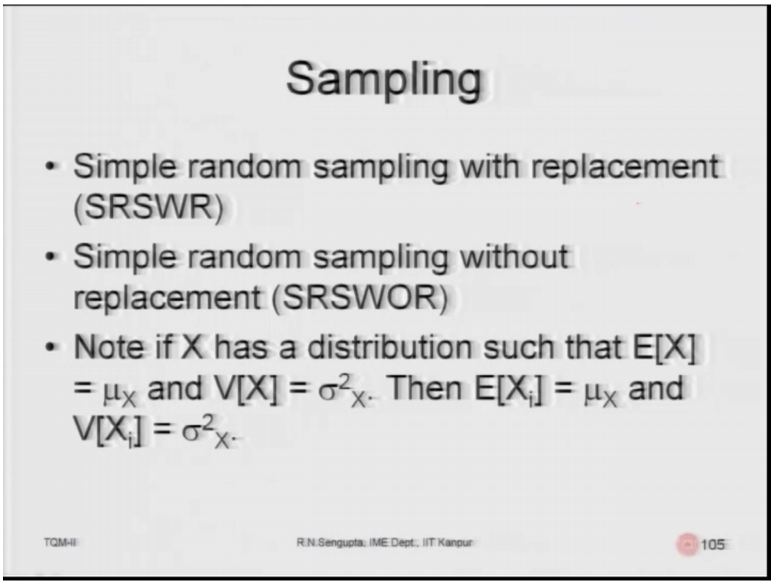
So, you want to basically find out in some cases the sample distribution would be given in some cases the sampling distribution as you found out. So, if you would ask that what is the sample distribution; you want to find, find out that it can be either related to the parameter which you are trying to find out from the sample distribution or it can be say for example, the random variable which basically knows the sample distribution.

So, in case if x is the random variable then from the sample you get the simple sample mean. So, if distribution of this is known or main task would be to find out the distribution of x n mark, or else it can be I want to find out see for example, x some star whatever it is. So, x star would basically be the minimum values between x 1 to x n I want to find out the distribution of that, so that is also important. So, that may be the sample distribution which I am looking for.

So, in the continuing with this example consider a population having the elements 1, 3, 6 7, and here n is 4 and the mu value would be basically the sum of all these 4 elements 7 plus 6 is 13, 14, 1, you add 1 and plus 3, 17, 17 by 4 is the population if we take n is equal to 2; that means, we take 2 observations at each go and then the 2 observations can be any combination can be 1 1 it can be 1 3 1 6 1 7, and the other extreme it can be 7 1 7 3 7 6 7 7.

So, then the possible values of the sample average can be 1 plus 1 divided by 2 which is 1 and the other extreme is basically seven plus seven divided by 2 which is 7. So, this would be the sample statistics before which we need to find out the distribution function.

(Refer Slide Time: 14:12)



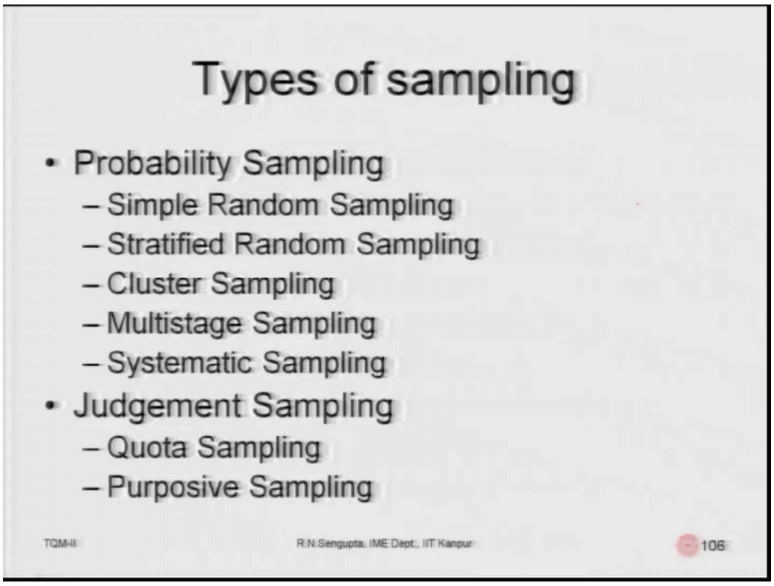
Now, simple random sampling can be with replacement and without replacement. So, with replacement means that you pick up one observation and you do the checking find out what is the observation and then you do continue doing this random sampling with replacement.

So, you pick up check put it pick up check put it and without replacement is basically pick up find it out the find out what is the probability and then you remove that object and then basically do the sampling or find out the probability again, but the problem is here. So, say for example, you have a big sample of size 100 there are chits mark 1 to 100 each being only happening once. So, if you do it with replacement the probability corresponding to pick up any chit remains the same.

So, if you pick up 1 finally, its probability as 1 by 100 noted down as 1 by 100 and put it in the box. If you continue doing it the probability always remains the 1 by 100, but in other case if you pick it up and basically remove it so; obviously, it means the probability of picking up 1 the next time if 1 has been pulled picked up in the first time basically becomes 0, or say for example, the probability of picking up number 2.

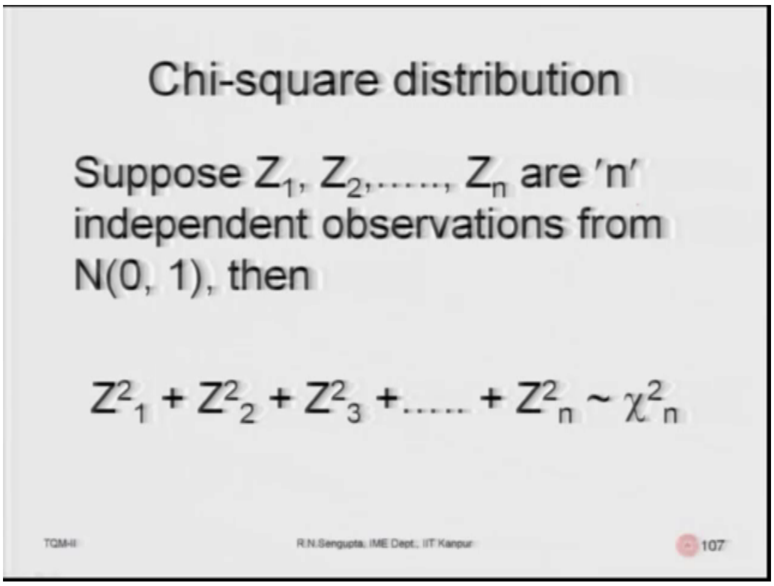
If number 2 has not been picked up in the first trial it will be 1 by 90 a 1 so; obviously, the probability will change, but it should be remembered, but if this actual population of the sample size is huge and then the corresponding differences in the probabilities doesn't change much which would not be much of a difference in the actual answer, note if x as a distribution such that e of x is mew x and variance of x is sigma square then the expected value of X i and variance of X i continue to remain as mu suffix x and sigma square suffix x.

(Refer Slide Time: 16:01)



Now, there are different types of sampling, so they can be probably sampling under which you can have simple random sampling, stratified sampling, cluster sampling, multistage sampling, systematic sampling, and the judge judgmental sampling, you have quota sampling purposive sampling and all these things.

(Refer Slide Time: 16:19)

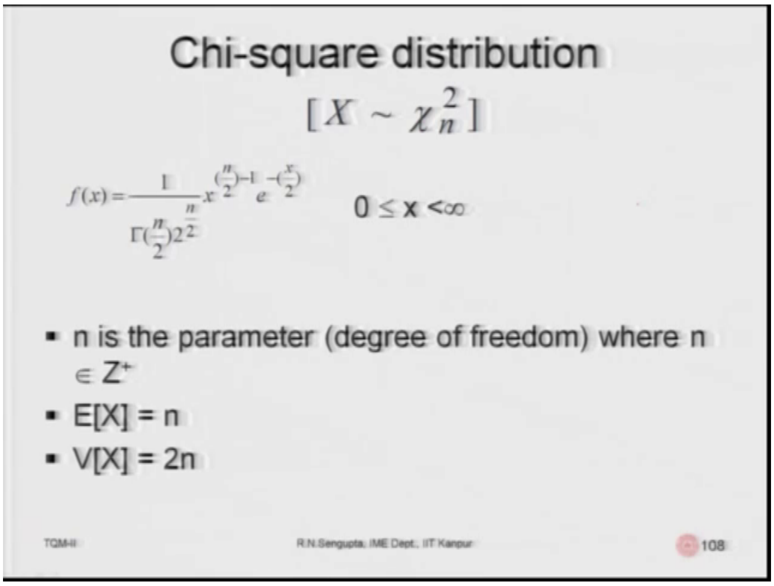


Consider the chi square distribution, so if you have a basically chi square distribution how it is formed. So, consider you have n number of Z distribution Z is basically the normal distributions who consider that a boxes each are z distributed and these boxes has infinite number of observations.

So, you pick up 1 observation that would be the Z distribution you pick up each observation square them add them up and the value which you get you keep aside, you do the second time pick up the second observation from box 1, second observation box 2, nth observation box n square each of them individually add them up and write down the values.

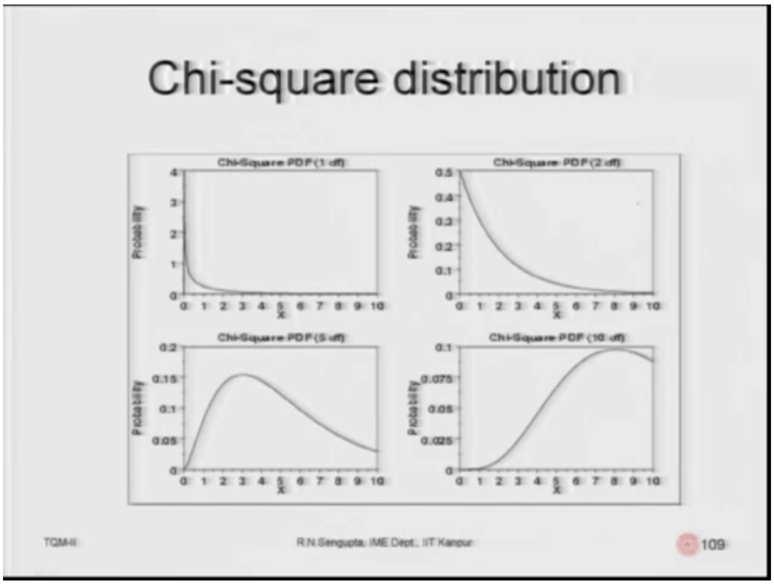
So, if you continue doing it the values which you have the square sum of the squared values if you plot them then the distribution basically is chi square. So, this is what we mean by sky square distribution. So, we coming back to this suppose Z 1 to Z n are n independent observations from normal 0 1 then Z 1 square. So, 1 means basically the suffix Z 2 square the 2 is basically the suffix till Z n square are chi square distributed with degrees of freedom n.

(Refer Slide Time: 17:27)



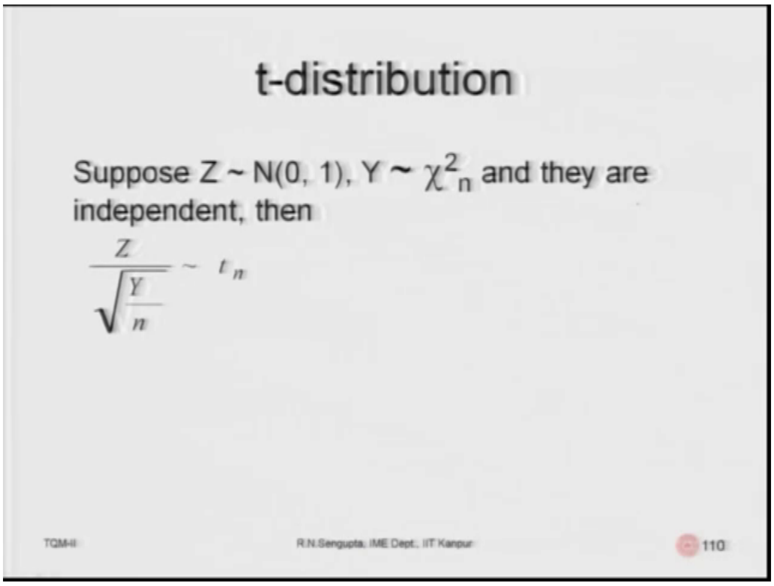
So, this is the distribution of chi square and the n if n is the parameter which is the degrees of freedom. So, the expected value is given by n and the variance is given by 2 n.

(Refer Slide Time: 17:38)



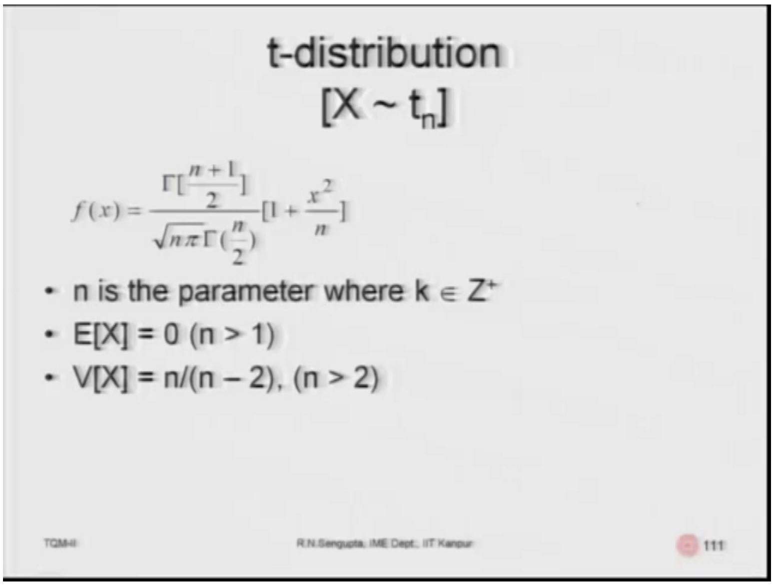
And the chi square distribution looks like as given here, so depending on the degrees of freedom whether 1 or 2 or 5 or ten the distribution would basically take the shape as given. So, along the x axis you have the sky square values. So, those are not the p d fs and along the y axis you have the p d fs correspond in the chi square.

(Refer Slide Time: 18:00)



Now, you have the t distribution, so how do you form the distribution consider that the y is already chi square and then degrees of freedom you have another distribution which is Z which is standard normal then if you find out the ratio of Z divided by the square root of chi square by its degree of freedom you will have basically the t distribution. So, that will give you one thing should be remembered the t distribution is almost symmetric to the normal distribution. So, if we increase the degree or the degrees of freedom or the n value then t distribution becomes exactly equal to the normal distribution in the sample size increases.

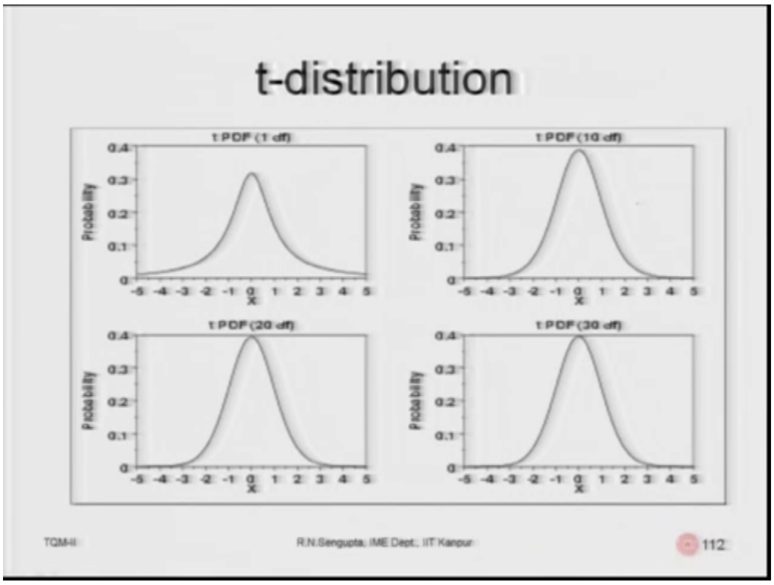
(Refer Slide Time: 18:40)



So, t distribution p d f is given by this expected value is 0 because t distribution as I said is looks exactly symmetric to the Z distribution and the variance is given by n by n minus 2.

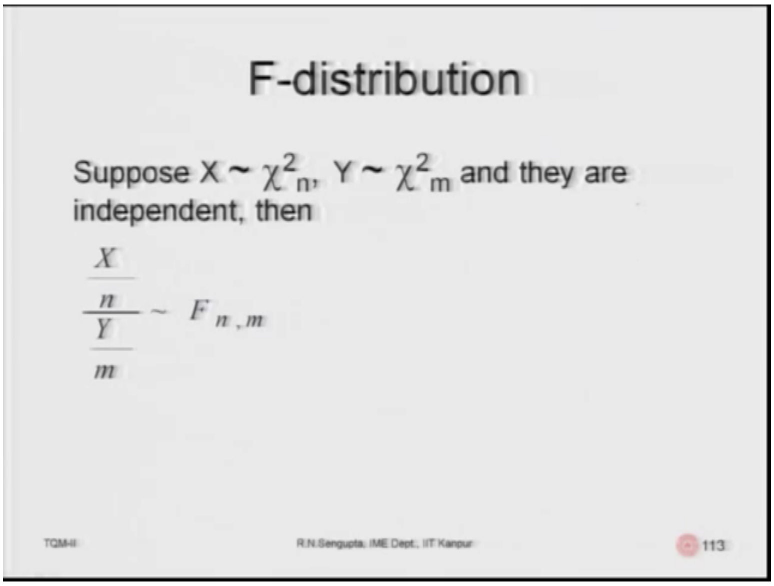
So, you can find out and; obviously, it would mean it is very interestingly this as n increases the ratio becomes 1 which is the variance of the standard deviation of the standard normal distribution.

(Refer Slide Time: 19:04)



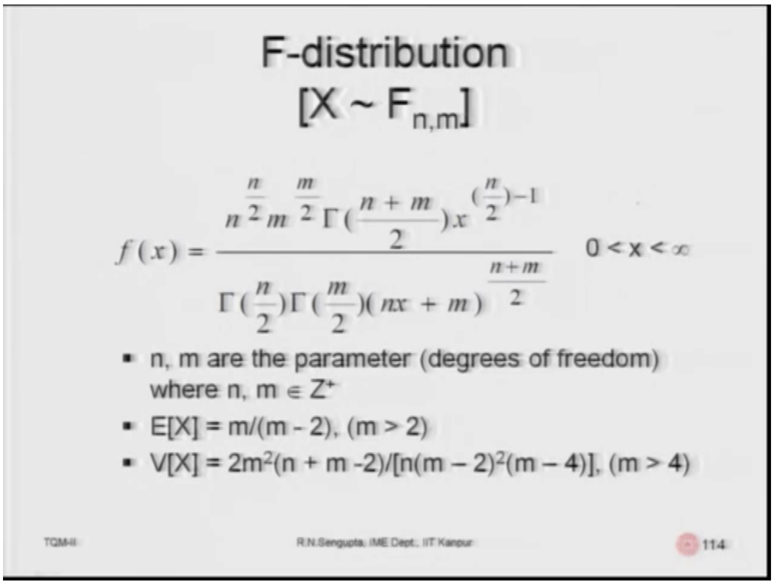
So, if you look at this pics set up values or pictures for degrees of freedom of 1, 10, 20, 30, the p d f of the ts distribution is drawn along the y axis e of the p d f along the x axis you have the t values and if you plot them it becomes almost exactly equal to the normal distribution and the degrees of freedom increase.

(Refer Slide Time: 19:25)



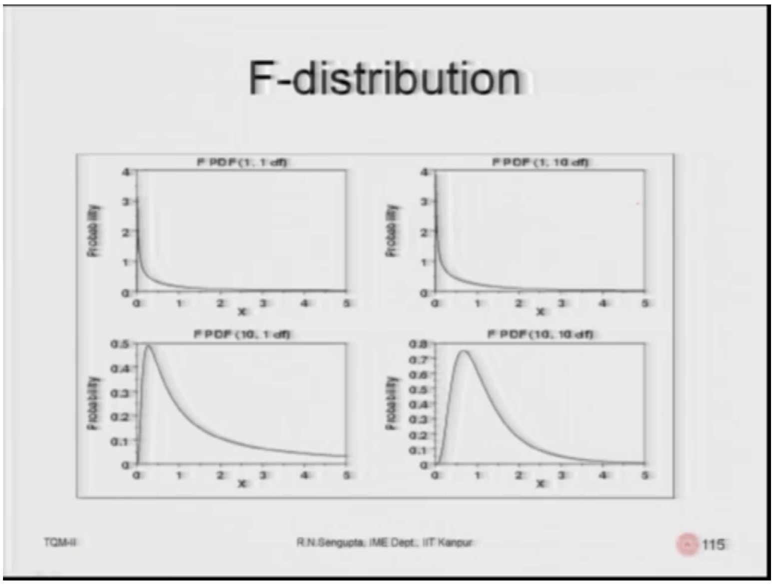
The F-distribution is basically a combination of 2 chi square. So, consider x is pi chi square with degrees of freedom of n and y is chi square with degrees of freedom m. So, if we find out the ratio of chi square divided by its degrees of freedom n and that is divided by chi square with divided by degrees of freedom which is m then that ratio is known as F-distribution with degrees of freedom n comma m if you revert the ratio; that means, Y by m divided by X by n will give you basically the F-distribution with degrees of freedom m comma n.

(Refer Slide Time: 20:07)



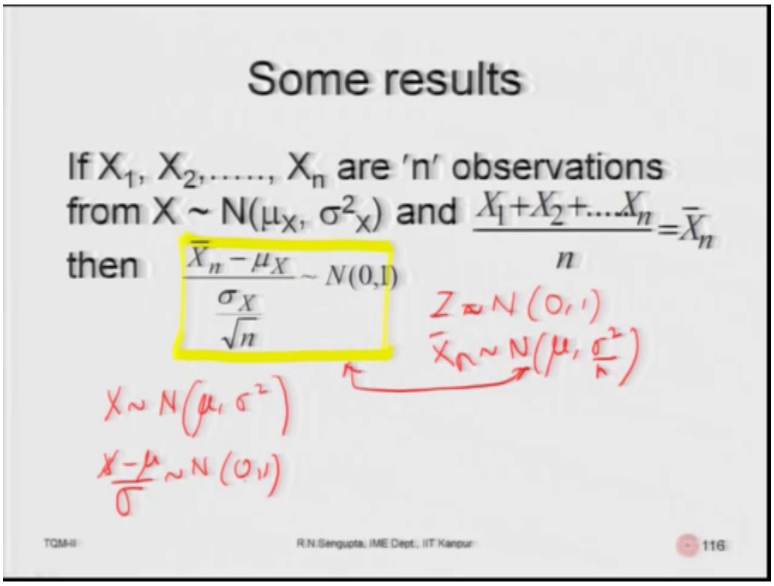
So, the F-distributions p d f value is given where n m n n m n are the parameters expected values m divided by m minus 2, or m divided by n minus 2 depending on which ratios you are taking and the variations would be variance would be given; as it is given where if it is n by m or m by n the values will be taken accordingly whether n or m you can decide it accordingly depending on the ratio.

(Refer Slide Time: 20:34)



The F-distribution is also not symmetric like I square. So, depending on the degrees of freedom of 1 1 because now they are 2 degrees of freedom m and n depending on 1 1 or 1 10 10 1 and 10 10 10 the distribution would be given as shown in front of you.

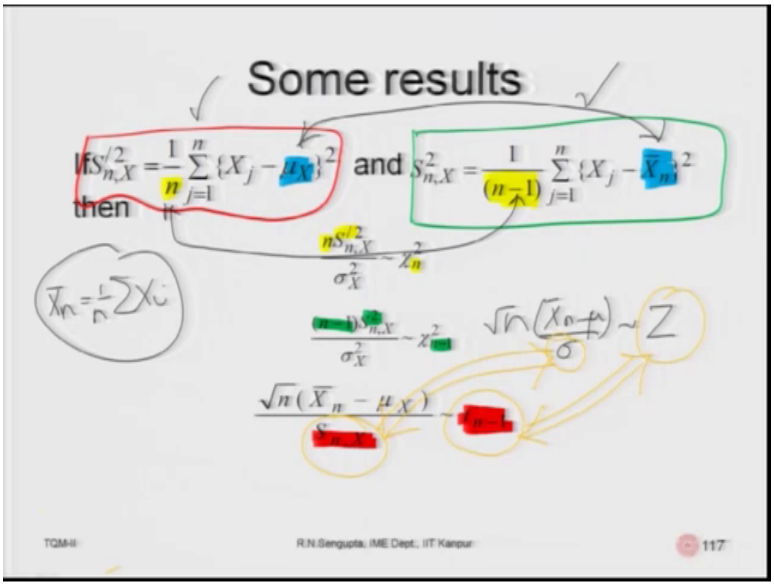
(Refer Slide Time: 20:51)



So, some interesting results are if X 1 and X 2 till X n are n observation from the normal mu suffix x and sigma square and the sum which is X 1 till X and divided by n is basically given by X bar then; obviously, we will know which we have discussed in details just before just at the beginning of few slides of this picture that this value is the z distribution.

So, this is a distribution because Z is normal 0 1 and this x bar n normal mu sigma square n and based on that you are getting this standard normal in case x is normal with mu sigma square then; obviously, it will be x minus mu by sigma that is do a distributed n 0 1 and you do your calculations accordingly.

(Refer Slide Time: 22:04)



So, some interesting results are considered you have a sample and you want to find out the sample standard deviation which is also known as standard error or you want to find out the sample variance. So, sample variance can be found out in two different instances, number 1 when the population mean is known and 1 case is when the population mean is not known.

So, when the population mean is and given then; obviously, the standard error whole square which is the variance of the sample is given by 1 by n summation of X i minus mu whole square that summed up, but if you are you do not have any information of the population mean. So, what you do is that you would use that sample set of observations for the first time and find out the sample mean.

So, mu is replaced by X bar n, so the moment you use that for the first time you lose 1 degrees of freedom hence it is divided by 1 minus divided by n minus 1 as highlighted.

So, I will try to highlight it with different colours. So, what is important to note down I will just highlight it once again for the benefit of the student.

So, this is what is this part is important along with this and another part which is important is this along with this. So, as, so this gets replaced here and as you utilise in this you lose 1 degree of freedom here, now if you come back use this concepts and also remember these.

So, utilizing these two concepts this one for the mean and this one for the standard error you will find out the following results, which is which are as follows the ratio of the staff the square of the standard error or the sample variance divided by the population variance multiplied by the n factor would give you a chi square with n degrees of freedom.

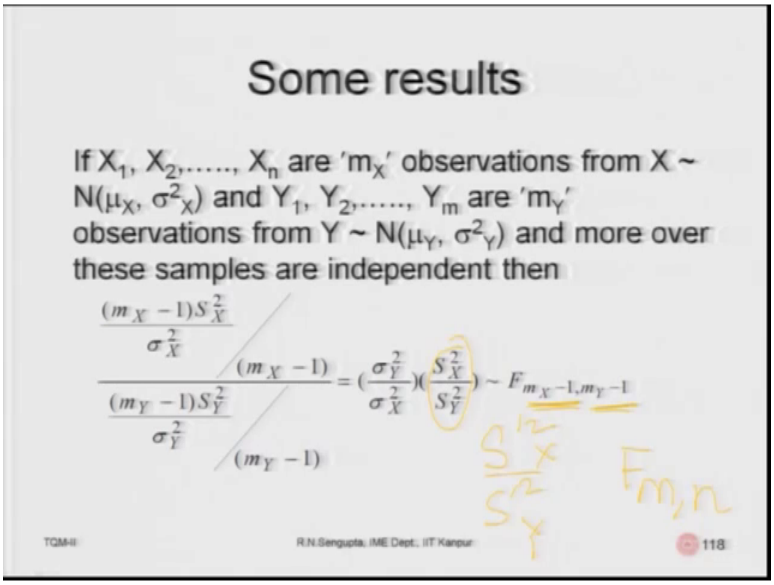
So, remember this I am I will again highlighted from few things with yellow colour. So, if it is dash this is n; this is n in the other case if it is without the dash this n minus 1 this is n minus 1. So, this would make things clear to you where the changes are and they are very logical; obviously, the proof is not very intense it can be proved, but it is not actually necessary for this TQM force.

So, this will give you the chi squared with n degrees of freedom chi square with n minus 1 degrees of freedom, and in the other case when you have the when you want to find out the distribution corresponding to t t n minus 1 so; obviously, in that case that sigma square which is the population variance is being replaced by the standard error of the sample and as you replace the standard error in the sample you will get the t distribution with n minus 1 degrees of freedom so ok. So, let me again say why I did say, but I didn't highlight.

So, I use another colour or a red 1 yeah, so this is being replaced in place of sigma square. So, and; obviously, this becomes I am not gonna highlight this remains square root of n because it is coming down. So, in this case it is t n minus 1 t d t distribution and if we compare this.

So, this was what in that case it would have been x n. So, the changes where they are occurring and just use another colour, so this and this are important to note and this and this are other important things to note.

(Refer Slide Time: 26:55)



So, if X 1 to X n r and r m x observations from, from distribution normal with mu and sigma square as the mean and variance and Y 1 2 i m are another normal distribution with mu and sigma square with this a corresponding suffix, then you can find out that the ratios of the corresponding distributions would be F-distribution, but there are two important points.

So, I want to mention number one in case these are x s without the suffix then in this case the f distribution loses 1 degrees of freedom both for m and n, in case they are the ratios of f s dashes in place of s. So, these are there, so in that case it will be f distribution of m and n. So, with this I will close this fourth lecture and continue discussing in the fifth lecture correspondingly the other concepts of inference techniques have a nice day.

Thank you very much.