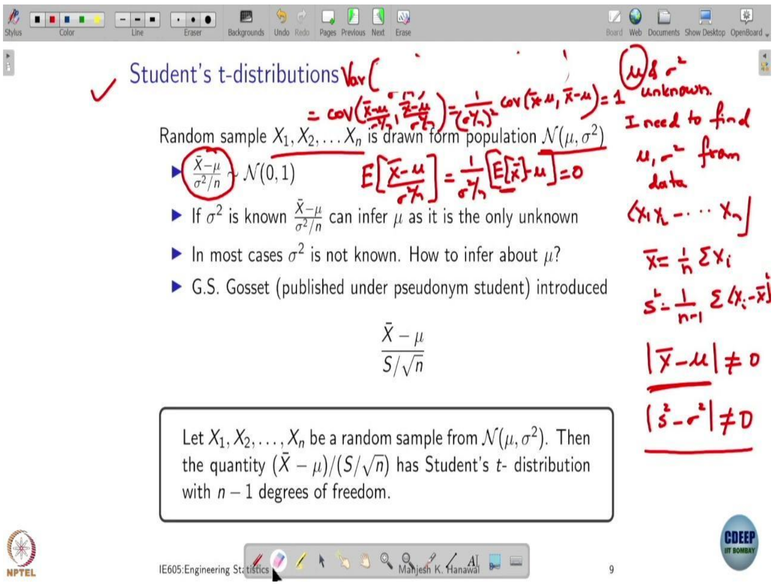
Engineering Statistics
Professor Manjesh Hanawal
Industrial Engineering and Operational Research
Indian Institute of Technology, Bombay
Lecture - 23
Student's t-distribution

(Refer Slide Time: 0:21)



Now let us say you have this random sample X1 to Xn drawn from this Gaussian mu sigma square. If you look into this quantity X bar minus mu divided by sigma square 1 the claim is this is going to be Gaussian with mean 0 and variance 1 or this is going to be a normal distribution. Everybody agree with this?

Student: Yes

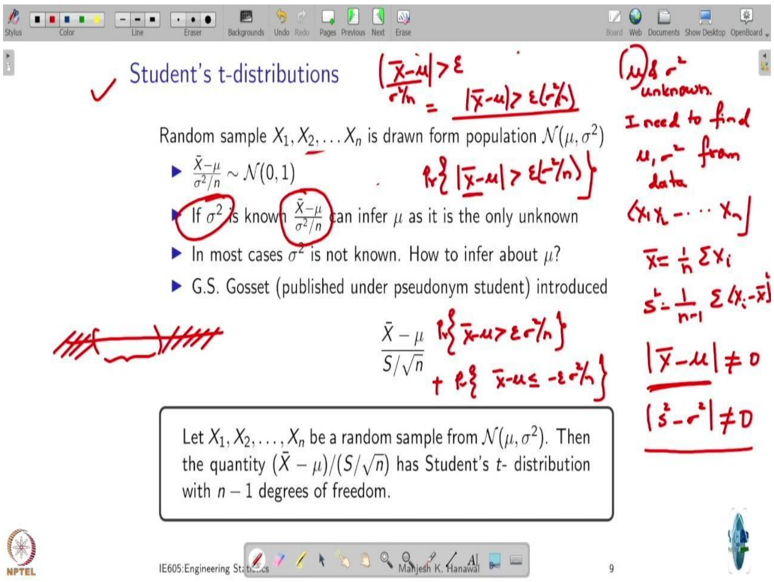
Professor: How to compute again?

Student: MGF

Professor: Again, go and do your momentum generating function calculate for this and you will actually see that they will end up getting. One quick way to verify that this is indeed have a mean is like how to, what is the, if you compute this. Can I write it as like this because sigma squared by n constant, I pulled it out and expectation of X bar minus mu and I know that expectation of X bar is mu that is why it is 0. And in a similar way you can also compute variance of X bar minus mu by sigma square by n which is nothing but covariance of X bar mu, agree with this?

And now, we know that this is nothing but 1 upon sigma square by n whole square covariance of X bar minus mu X bar minus mu. Now, if you just expand this you will see that this will actually turn out to be just 1.

(Refer Slide Time: 2:51)



Now first let us simplify further break down the problem. Instead of assuming that both mu and sigma squared unknown let us assume that sigma square is known, what is unknown is only the mean value is unknown. And now if that is the case can I understand from this value how much is the difference? Now, only mu is unknown and I want to quantify how much is the difference. I want to basically let us say I want to ask this question whether this is going to be greater than epsilon if I want to ask this question.

Now, what I know is instead of this what I will consider is I will consider this quantity which is nothing but X bar minus mu greater than epsilon sigma square by n. Now, can I ask what is the probability that X bar minus mu is greater than epsilon sigma square by n. Can I compute this probability?

Student: Yes

Professor: What is that phi of.

Student: 1 minus 4.

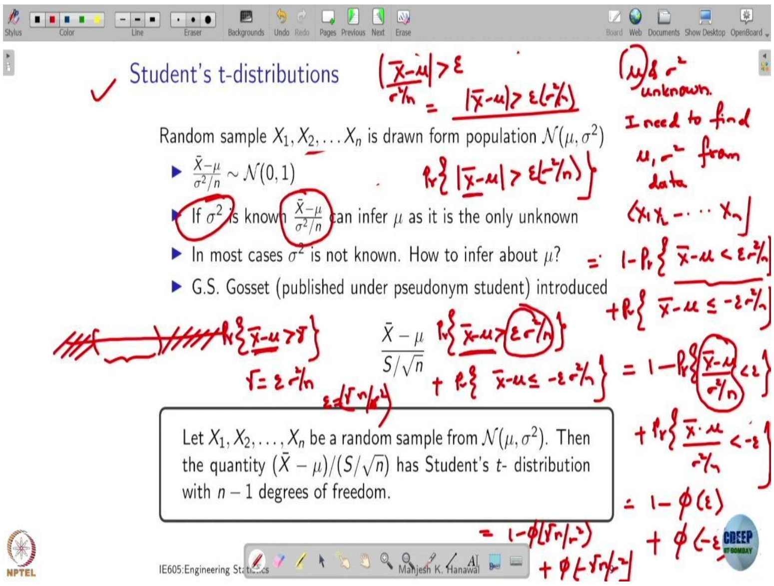
Professor: X bar, what is the distribution of X bar minus mu?

Student: Normal zero comma sigma square by n.

Professor: Everybody agree? Now, I cannot directly use the phi function here. I need to do couple of more steps here. What are those couple of more steps? Let us write here this is nothing but probability that X bar minus mu greater than epsilon sigma square by n and no wait a minute, so, what I want is this difference I want it to be lying in this region outside. This difference I want it to be exceeding some region let us write this.

This X minus mu should be maybe you should write like this or this and this, operated at X bar minus mu greater than epsilon sigma square by n plus probability that X bar minus mu less than minus epsilon sigma square by n, is this correct? I am just writing this part into these two probabilities like when this is positive quantity, it should be there. When it is negative, it should be equal to this.

(Refer Slide Time: 7:20)



Now, what I know about this? I can further simplify this step, where I can write, now this is nothing but 1 minus probability that X bar minus mu less than epsilon sigma square by n plus probability that X minus mu less than or equals to minus epsilon sigma squared by n and by our definition, what is this quantity? But notice that X minus mu what is the distribution of X minus mu and when I want to apply phi function, I want to make sure that the distribution of this quantity is normally distributed. So, what I will do is I think I should have avoided this circus, what should I be done is I should have kept is X minus mu sigma square by n is less than epsilon and plus probability that X minus mu sigma square by n less than minus epsilon.

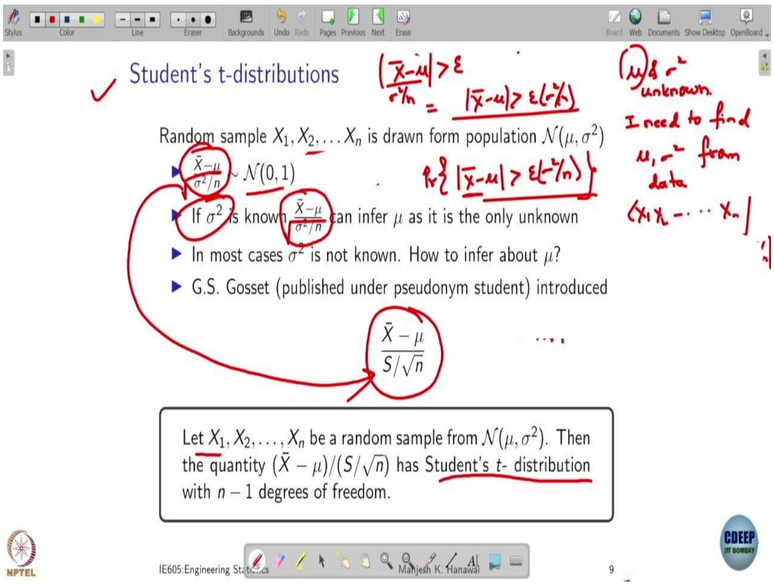
And now, I know that this quantity is what? Normally distributed. So, this is nothing but 1 minus phi of epsilon. And what is this other quantity? This is phi of minus epsilon. Now, notice that this all I could do provided my sigma square is known. Now, I know that if I know that now I am able to understand how much is my difference. Actually there is one small caveat here like the way I used I am now basically asking my difference X bar minus mu to be greater than this quantity, now I am but this can be anything that is given to you.

Suppose let us say you have been asked X bar minus mu let us say you want to find it this value to be greater than gamma. Now all this method works if I set my gamma to be epsilon sigma square by n and so actual sigma is gamma n by sigma square. So, if I want to know that my difference is greater than gamma, then I should be writing this epsilon to be gamma n divided by sigma squared. So, this should be actually 1 minus gamma n by sigma square plus gamma minus gamma n by sigma squared. Is that clear? It is clear, you wanted me to do more calculations? Is that fine, all these details?

Now, but unfortunately, many times we may not even know gamma square. Both mean and variance could be unknown. But in that case, I cannot apply this directly. Then what is the method for that? For that we have a method which were proposed by one of the famous statistician, long back, I think, in early 1800s, GS Gosset, I do not know it is 1800s or early 1900s, long back. Statistics is a very old subject. So, lot of people have worked in about lot of actually, development in statistics happened in early 1900s. So, lot of these results are pretty old actually.

But now, if you see that most of these results are now repackaged. And now you see them as machine learning methods. When you do machine learning, actually, you will be basically doing this. You will have in data trying to extract information from the data.

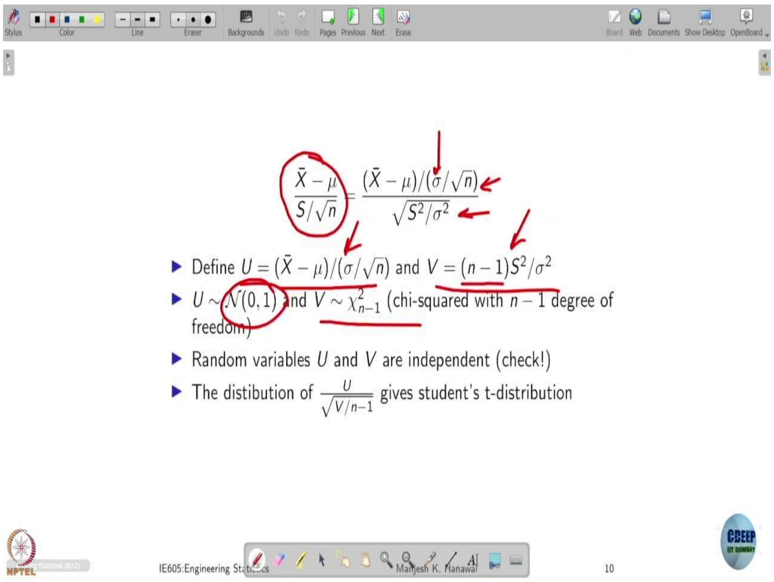
(Refer Slide Time: 12:21)

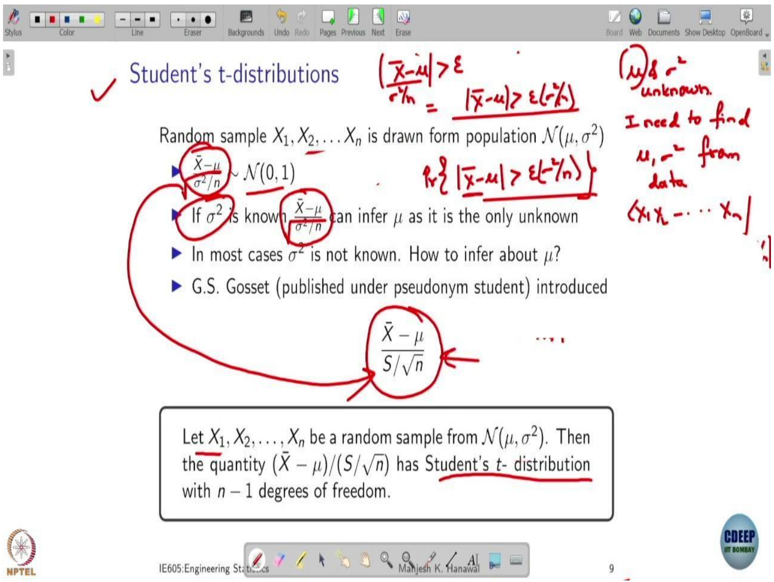


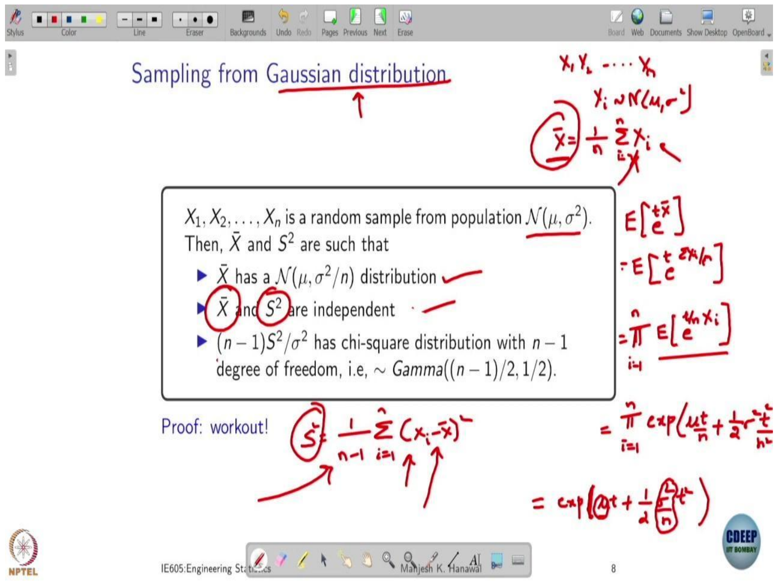
Now, when sigma squared is known, you could do this and you got a normal distribution. And that helped you to find out what is the error? Now, but when you do not know, I have to do find some proxy for sigma square. What is a good proxy for sigma square? Sample variance. So, that is what I will do is instead of sigma square, I will go and put sample variance I think, why is this, did I make any mistake, should it be square root or it should be squared root everywhere. Everybody agree, it should have been square root of sigma square by n instead of square root by n.

Now, that is what now, sigma squares proxies S square. And when I put under square root, it becomes simply S by square root n. And now I am saying that, I do not know sigma square, let us take it sample estimator, and I will put it like this. Now, the resulting distribution of this, notice that X bar is a random variable, S is a random variable. So, this whole quantity is a random variable. When the samples are coming from population, which is Gaussian with mu sigma square than this quantity, X bar minus mu S by n has a distribution, which is called Student t distribution with n minus degrees of freedom. Now, what is that t distribution?

(Refer Slide Time: 14:15)





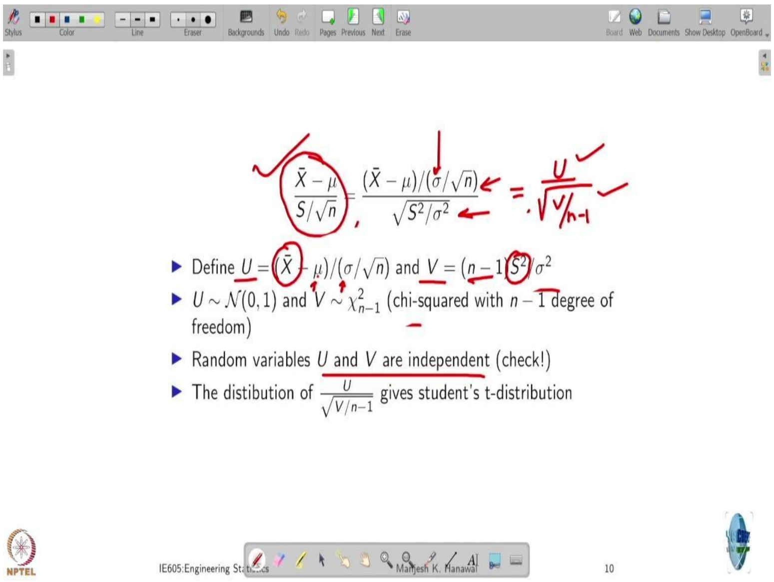


Now let us try to find out actually, what is the distribution of this quantity? And how does this t distribution look like? We will do some manipulation. This is the quantity of our interest X bar minus mu divided by S by square root n. What I will do is I will divide and multiply by sigma by square root n. If I do the sigma by square root n, you will get this expression. Notice that I do not know what is the actual value of sigma. But let us say that I could do this, sigma by square root n, I have just done this.

Now, I am going to look into this as a ratio of two random variables, the numerator is this quantity and the denominator is this quantity. What is the purpose of doing it like this separating numerator and denominator? If you now notice is the numerator is a quantity which I know which is X bar minus mu sigma by n square which has a Gaussian distribution and the denominator has this S square by sigma square suppose, if I multiply it by n minus 1 I know that we just discussed that this as a chi square distribution with n minus 1 degrees of freedom.

So, what I have basically done is and also notice that we just discussed that this quantity and this quantity they are independent of each other when it is Gaussian distributed, is not it? So, this quantity is just like this normalized version, centralized and normalized and this was sigma square, say square by sigma square when multiplied by n minus 1, we notice that this is a chi square distribution and these two are independent. So, you have to be careful like what we actually said is X squared and S squared are independent.

(Refer Slide Time: 16:45)



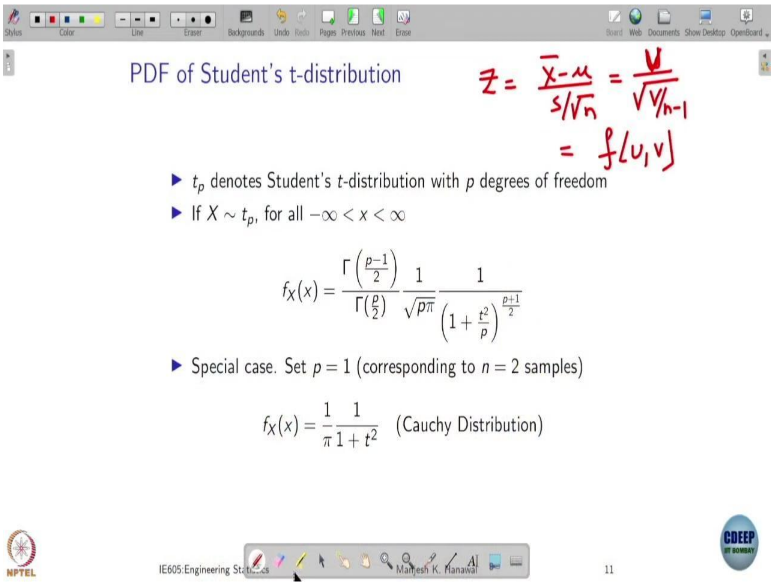
But here we are saying this quantity U and this quantity V, they are independent. This needs little bit thinking. We know that X bar and S square are independent. But here S squared is being subtracted by constant mu and normalized by sigma. Whereas this quantity S squared is also multiplied by constant n minus 1 by sigma square. So, both these S square here are multiplied by constant. And this X bar is also manipulated with some constant. If you just add some constant or multiply constants, does this change the independence nature? No. So, that is why we are doing this.

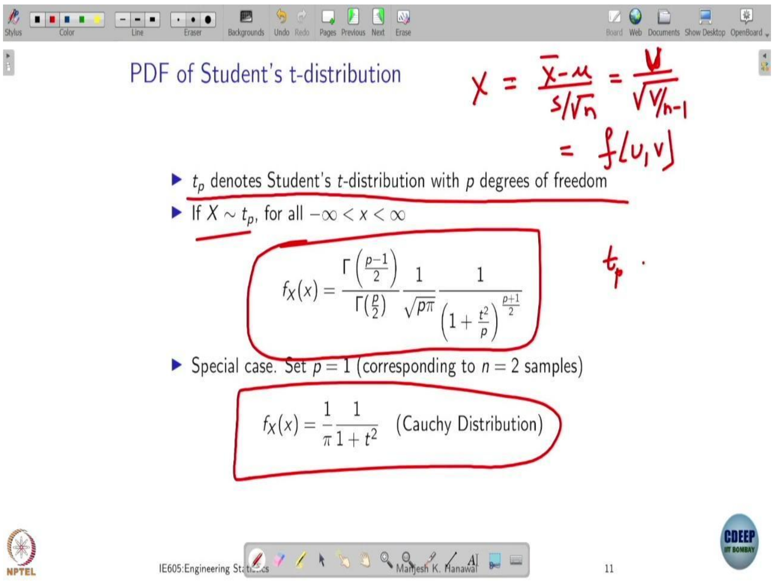
Student: Sir, here sigma is not known

Professor: Sigma is unknown, but you just take some sigma value. It is constant here. It is unknown but in this computation, it is a constant. We are notice that I have write an exactly sigma, I have not put it as an estimator or something. So, this is that constant only. Now, if you now look into this, this quantity here nothing but in terms of our representation, numerator is simply U and the denominator is simply V by n minus 1 and this quantity and the denominator they are independent of each other.

So, this helps us to compute the distribution of this quantity which we are calling as t distribution. So, this helps us to compute the distribution of this quantity which we are calling as t distribution.

(Refer Slide Time: 18:44)



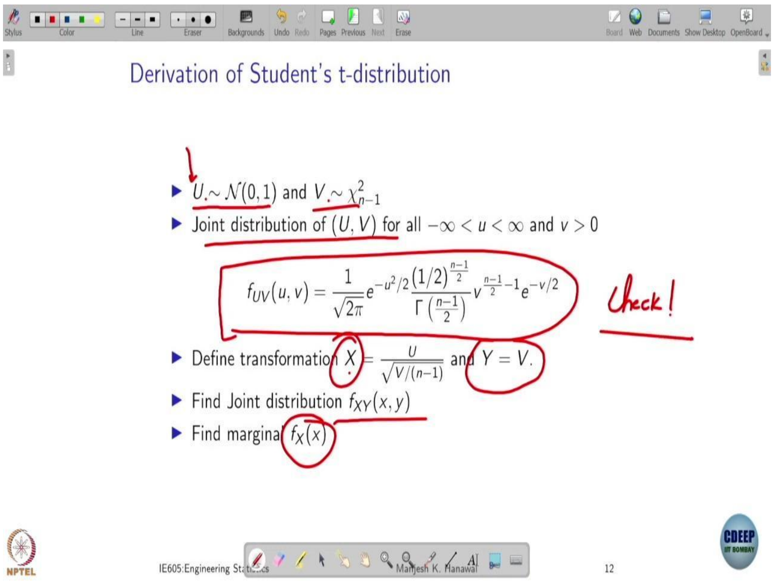


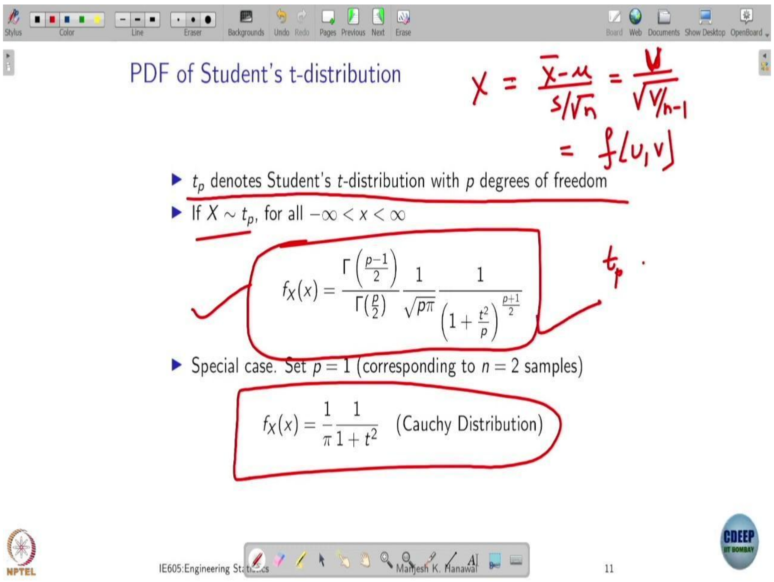
Now, if you use this property, now what we have let us call our Z. Z is our quantity of interest which is mu by S by square root of n, which now we are able to write it as U by V by n minus 1. So, I can treat as now Z as a function of these two random variables, U and V. And we know how to compute distribution of functions of random variables. If you recall all the methods that we studied earlier, where we use Jacobian and all, to find out the distribution of joint distributions, we can use that method here.

And now if you do that, I am jumping the step. Now, if you do that, you will actually end up the distribution of here I am calling this as X. This is X. This X will have t distribution denoted by t p. And it will end up with this kind of PDF function. And this is for all X between minus infinity to plus infinity. And this is exactly called t-distribution or denoted as t p with p degrees of freedom.

So, when I say now henceforth, my notation is when I write t p, this is a t distribution with p degrees of freedom and as a special case, when I said p equals to 1 this will end up with this simple form, which has a special name called Cauchy distribution. Now, quickly, you may be wondering how did this magically come up this distribution.

(Refer Slide Time: 20:45)





As I said this come from computing the joint distributions, so, let us take you have this U and V. And we know that U and V are independent and we know the distribution of U, U is Gaussian distributed and V is chi-squared distributed with n minus 1 degrees of freedom. Since U and V are independent, I should be able to write their joint PDF just take their product because of their independence. And now define two random variables X like this, Y like this, and then just to find the joint distribution of f of y in terms of that joint distribution of U and V that we have done using Jacobian method earlier.

And now, you will see that after that you find out after you find a joint distribution, you find a marginal that will exactly give you what I have written here. Again, this needs to be verified. That is why to have this need to check. Again, I will post you the book where these computations are done, but you need to go and verify that. It is very difficult and not really difficult. It will be like I do not want to sit out and work out all the details. This all the details you already know.