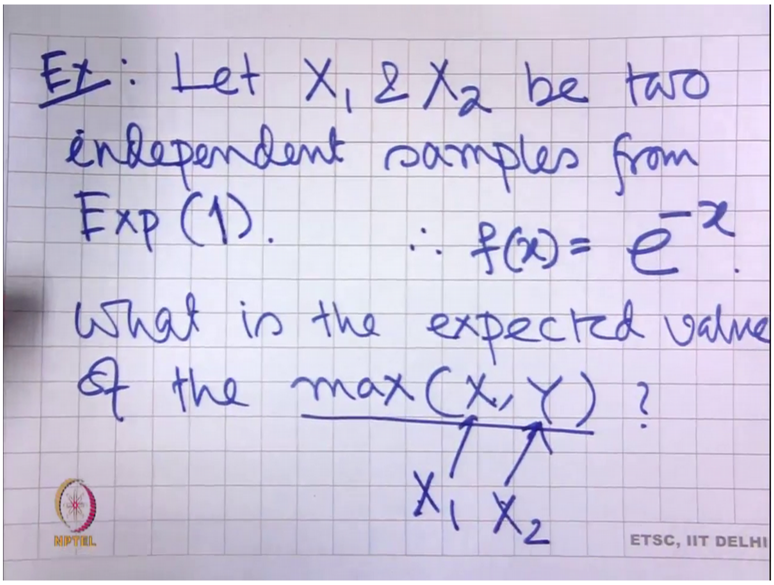
Statistical Inference
Prof. Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 11
Statistical Inference

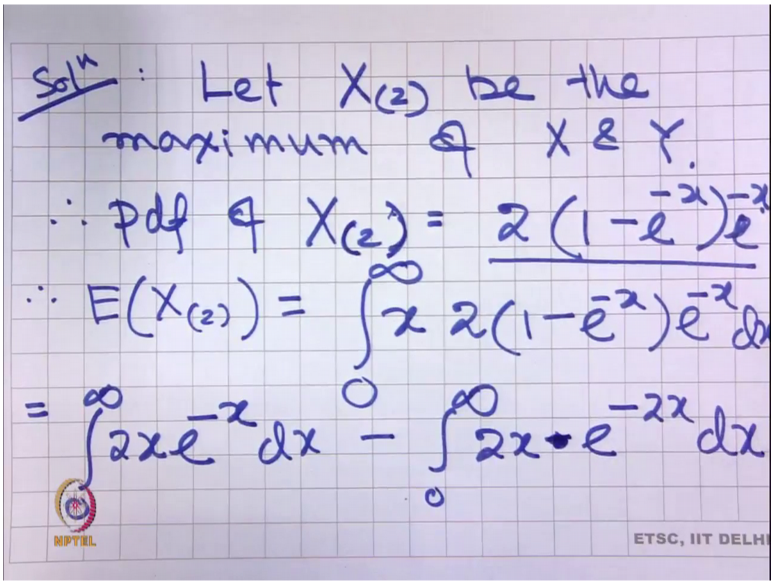
Welcome students, the MOOC's online course on Statistical Inference. This is lecture number 11, in this week we are studying Order Statistics, in the last two lectures we have developed the theory and we have seen some problems involving order statistics. In this class I will also solve quite a few problems which will help you in solving the problems given in the tutorial sheets and some more difficult problems in general.

(Refer Slide Time: 01:04)



Consider exponential distribution let X 1 and X 2 be, two independent samples from exponential 1. Therefore, f x is equal to e to the power minus x. What is the expected value, of the maximum of X and Y? Note that I have changed the notation from X 1, X 2 to X Y this is just for convenience.

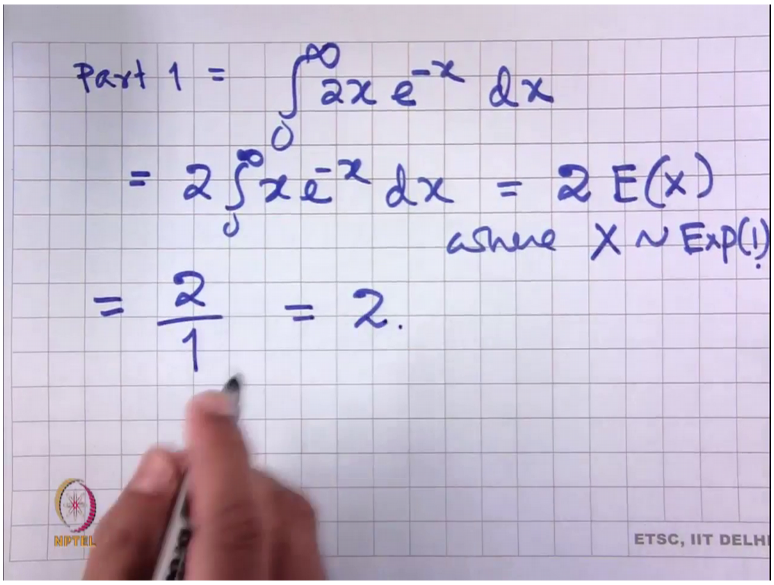
(Refer Slide Time: 02:31)



So, solution so what is the Pdf of X 2? Pdf of X 2 is 2 1 minus e to the power minus x. This is the capital f x or the cumulative distribution function and e to the power minus x is the small f x or the Pdf. Therefore, expected value of X 2 is equal to integration 0 to infinity x 2 1 minus e to the power minus x into e to the power minus x dx.

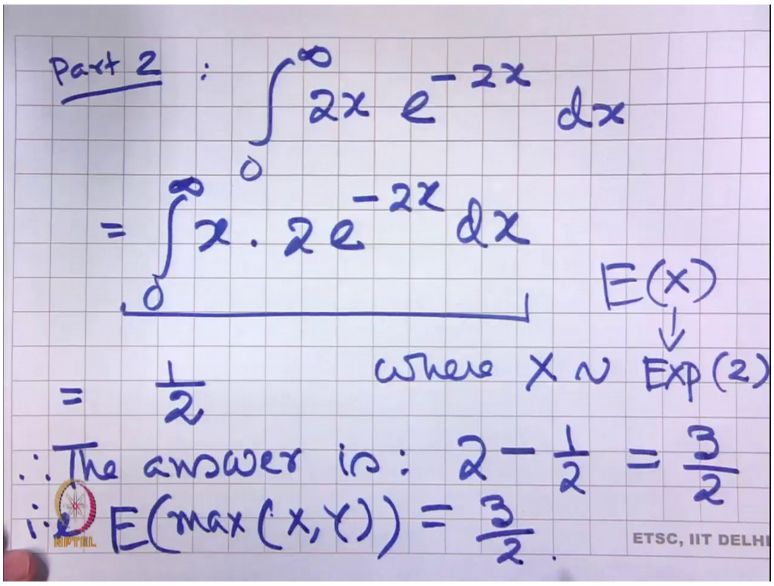
Now, we can split it into two parts, 0 to infinity 2 x e to the power minus x dx minus 0 to infinity 2 x minus e to the power 2 x into e to the power minus 2 x dx.

(Refer Slide Time: 04:41)



So, part 1 is equal to 0 to infinity 2 x e to the power minus x dx is equal to 2 into 0 to infinity x e to the power minus x dx is equal to 2 into expected value of x where x follows exponential with 1 and we know that if a random variable follows exponential with lambda. Then it is expectation is 1 upon lambda therefore, this comes out to be 2 upon 1 is equal to 2.

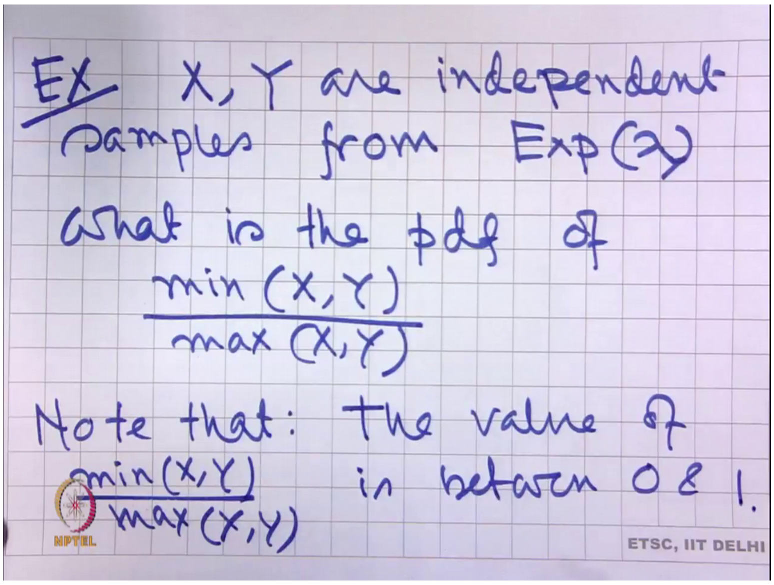
(Refer Slide Time: 05:37)



But 2 is 0 to infinity 2 x e to the power minus 2 x dx is equal to 0 to infinity x into 2 e to the power minus 2 x dx and this part is expected value of x where x is exponential with 2.

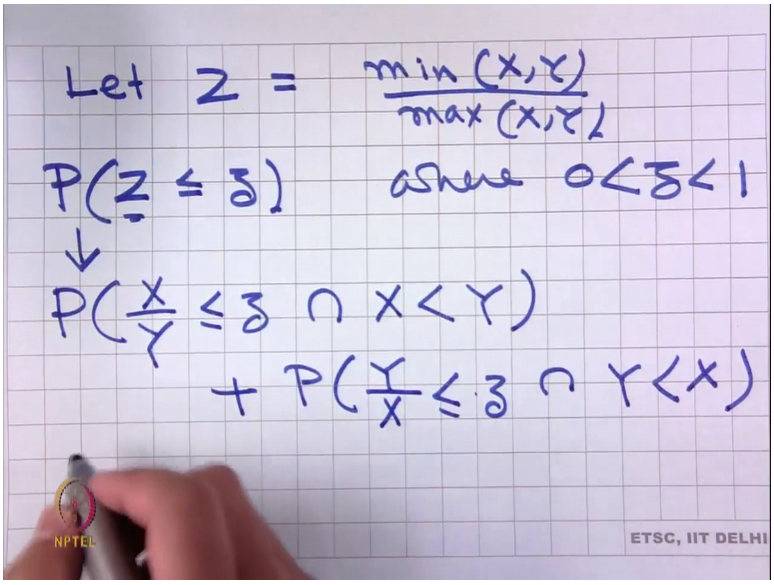
Therefore this expected value is going to be 1 upon 2 because if the parameter is lambda then expected value is 1 upon lambda. Therefore, the answer is 2 minus 1 upon 2 is equal to 3 by 2. That is expected value of max of x comma y is equal to 3 by two that is the answer.

(Refer Slide Time: 07:17)



Now, let me consider slightly more challenging problem the setup is same. We are working on X, Y there independent samples from exponential lambda. Our aim is to find out the pdf of minimum of X, Y upon maximum of X, Y. And note that these two are not quite independent. So, one way of doing it is we can look at that joint distribution of this to minimum of X, Y and maximum of X, Y, and from there we may like to compute the pdf of their ratios. I do it in a slightly tricky way. So, let us proceed note that this value is always between 0 and 1.

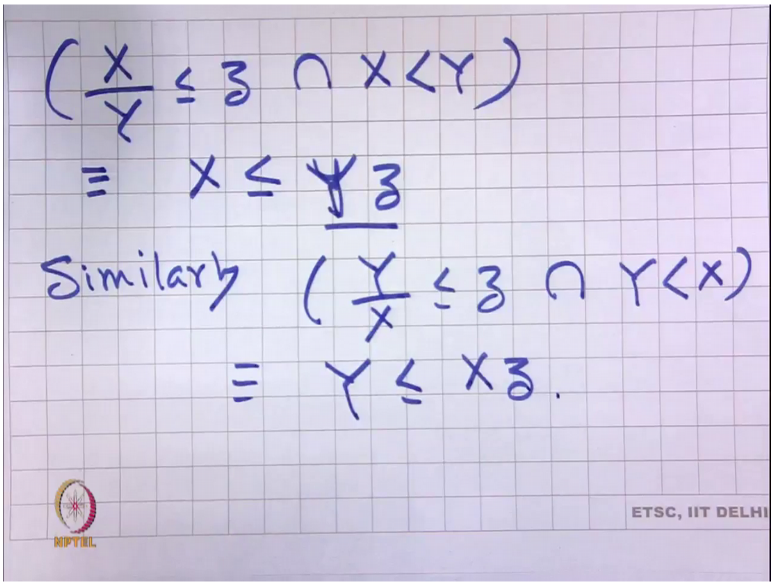
(Refer Slide Time: 09:17)



Let Z equal to minimum of X, Y upon maximum of X,Y. So, we are looking at probability Z less than equal to z that is we are looking at the cdf of Z. where 0 less than

Z less than 1. Now this event we can have in two possible ways, one is X by Y less than equal to z and X is less than Y plus probability Y by X is less than equal to z and Y less than X.

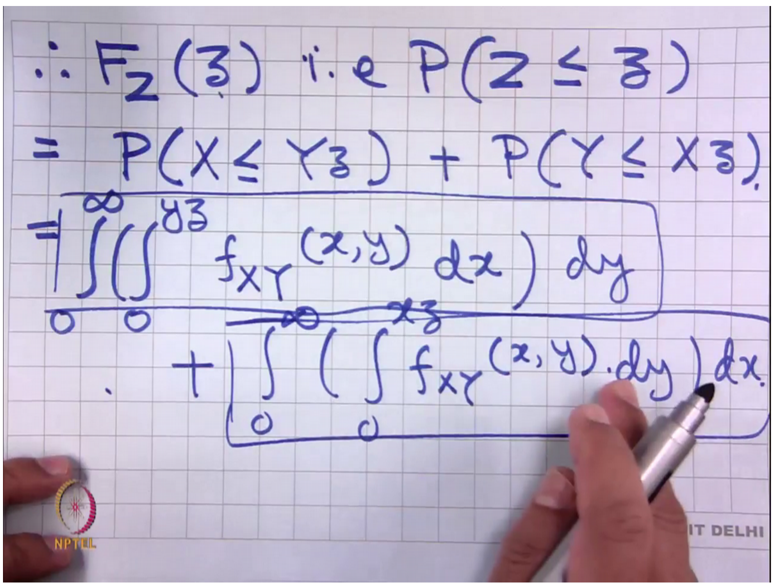
(Refer Slide Time: 10:41)



So, consider X by Y less than equal to z and X less than Y. This is equivalent to X less than equal to Y z and since z is less than one this is automatically satisfied.

Similarly, a event y by x less than equal to z and y less than x is equivalent to, y less than equal to x z.

(Refer Slide Time: 11:35)

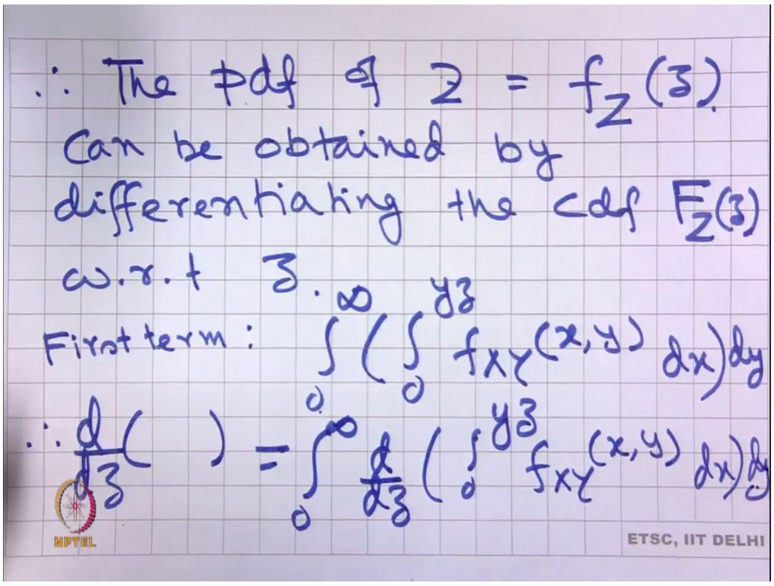


Therefore F z at z that is probability Z less than equal to z is equal to probability X less than equal to Y z plus probability Y less than equal to X z. Is equal to, In the first case; x will have to be less than equal to Y z and Y can be anything between 0 to infinity.

Therefore, range of y is 0 to infinity, but for a given y x can go only from 0 to y z. And here we are writing the joint density of x, y dx dy plus a very symmetric term which is 0 to infinity integration 0 to x z f xy xy dy dx.

So, this is clear this is the joint density of x, y both are independent. So, you can write it as a product of these two and then x for a given y x ranges between 0 to y z. Note that the z is fixed here and then y, I am integrating it over y from 0 to infinity and a very similar thing we are doing in the second case.

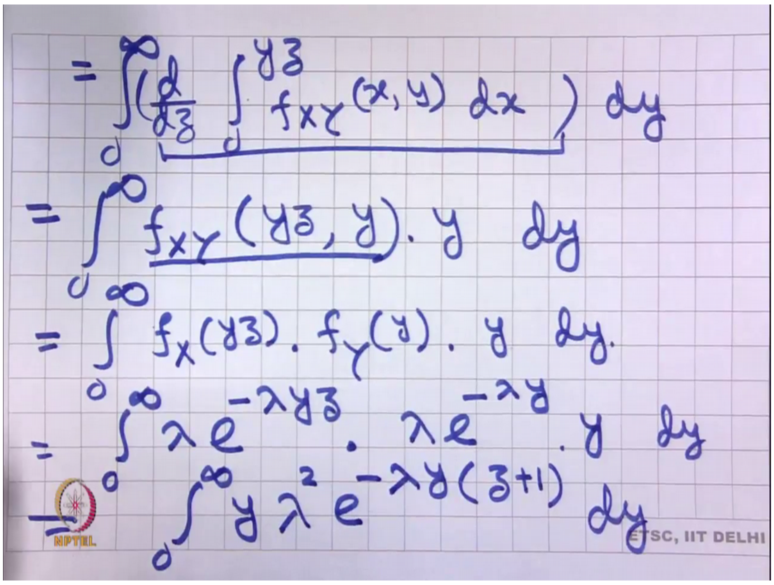
(Refer Slide Time: 13:53)



Therefore the pdf of z is equal to f Z z can be obtained by differentiating the cdf of z with respect to z.

So, let us consider the first term 0 to infinity 0 to y z, f of x, y x comma y dx dy. So, when we differentiate with respect to z look that the outer term outer integration does not involve any z. Therefore, dz of this term is equal to 0 to infinity 0 to y z f X, Y x, y dx dy is equal to.

(Refer Slide Time: 15:55)

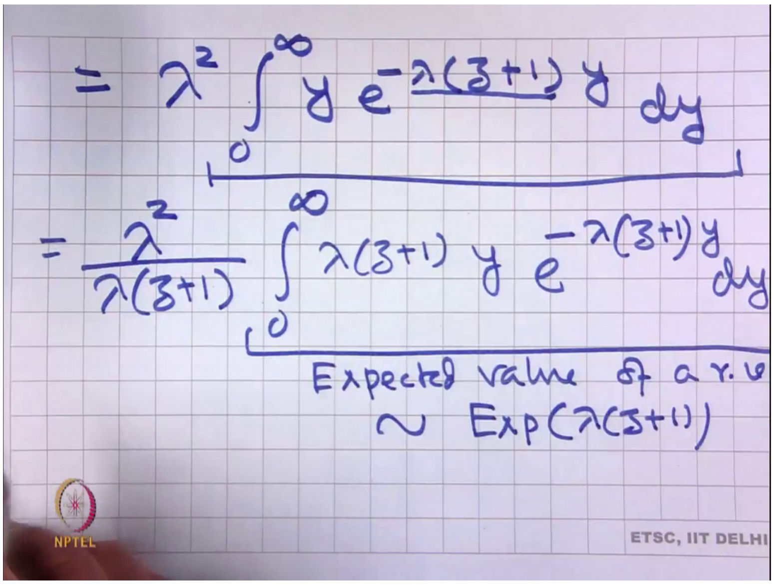


Let me write it again 0 to infinity d dz of integration 0 to y z f xy at xy dx the whole thing is dy is equal to 0 to infinity.

Now, we are cutting out this differentiation therefore, it is going to be f xy yz the value of x has to be replaced with y z, that is y z comma y multiplied by dyz dz which is is equal to y dy is equal to 0 to infinity what is fx y? Xis exponential with lambda and y is also exponential with lambda and they are independent.

So, we can write it as fx at yz multiplied by f y at y into y dy is equal to integration 0 to infinity lambda e to the power minus lambda y z multiplied by lambda e to the power minus lambda y y dy which is equal to integration 0 to infinity y lambda square e to the power minus lambda y into z plus 1 dy.

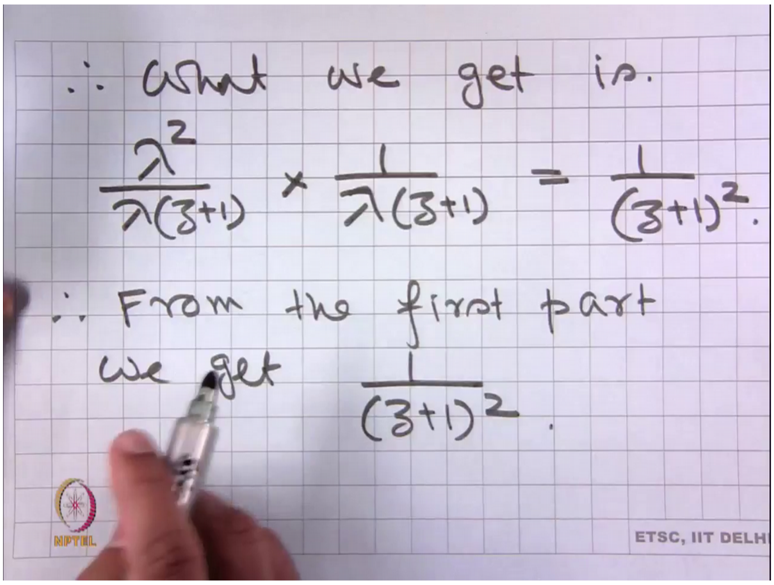
(Refer Slide Time: 18:25)



Is equal to lambda square integration 0 to infinity y e to the power minus lambda z plus 1 into y dy.

Now, let us look at this term, this resembles exponential distribution where for a fix z the parameter is lambda into z plus 1 therefore, suppose we write it as lambda square upon lambda into z plus 1 into 0 to infinity lambda into z plus 1 y e to the power minus lambda into z plus 1 y dy. Then basically this is expected value of a random variable which is distributed as exponential with lambda z plus 1. If you accept that part then we know that this expectation is going to be 1 upon lambda z plus 1.

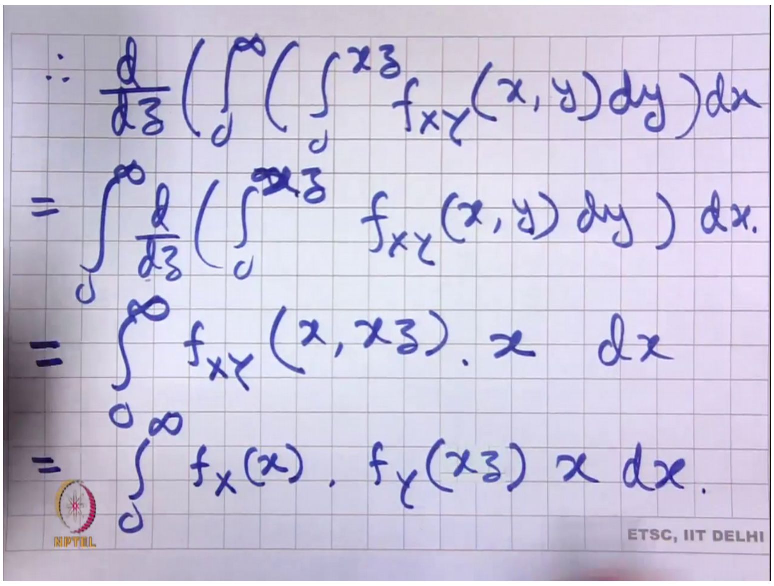
(Refer Slide Time: 20:05)



Therefore what we get is lambda square into lambda z plus 1, multiplied by 1 upon lambda z plus 1 is equal to 1 upon z plus 1 square. Therefore, from the first part we get 1 upon z plus 1 square. If we look at the holder expression we had this as the first term and this has the second term. And if we watch carefully we find that these two are basically the same thing because this is independent. So, in the second case when I am differentiating with respect to z it is going to be fx y at x comma x z and multiplied by x.

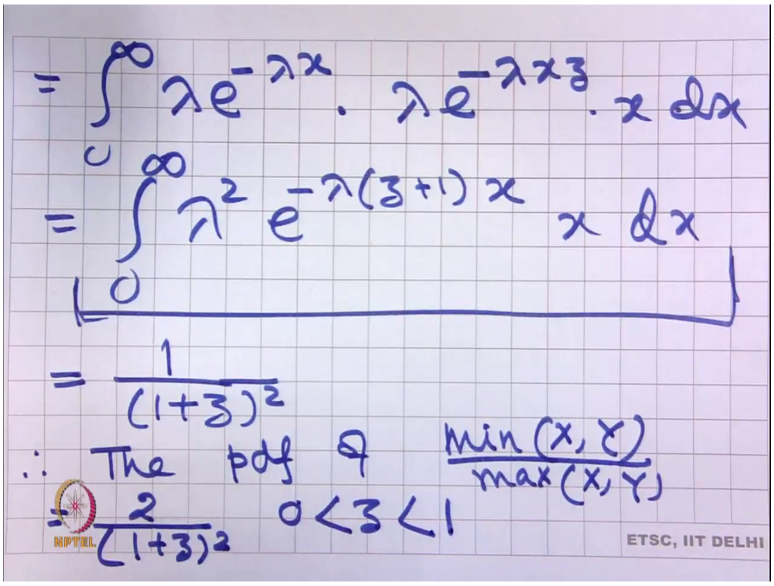
So, let me write it down d dz of 0 to infinity 0 to x z fxy x comma y dy dx. And in a very similar way what we did for that, first part this is going to be 0 to infinity d dz of 0 to infinity 0 to x z f xy x, y dy dx.

(Refer Slide Time: 21:45)



This is equal to 0 to infinity f xy at x coma. Now y is going to get this value x z, now we differentiate with xz with respect to z that gives us an x dx is equal to 0 to infinity fx at x if y at xz x dx is equal to 0 to infinity.

(Refer Slide Time: 23:09)

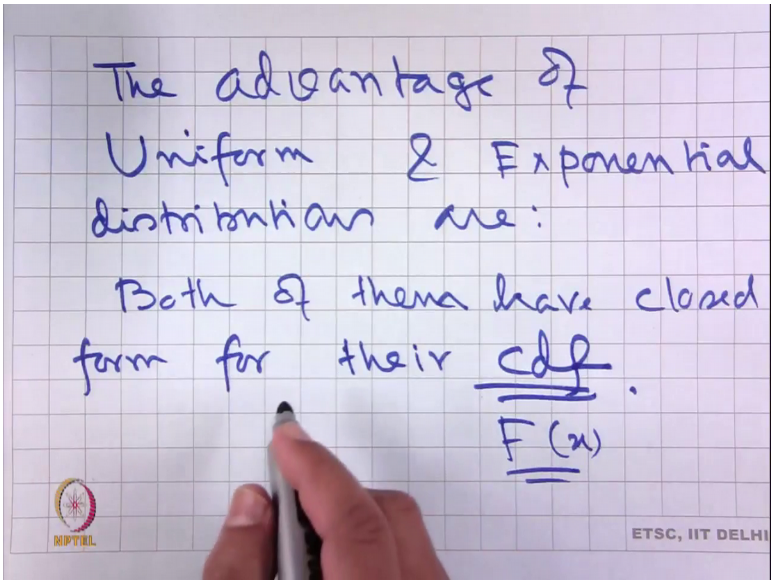


Lambda e to the power minus lambda x, multiplied by lambda e to the power minus lambda x z into x into dx is equal to 0 to infinity lambda square, e to the power minus lambda z plus 1 x x dx.

And now we see that this becomes same as what we get in the first term. Therefore, without evaluating it any further we can say that, this is going to give us 1 upon 1 plus z whole square. Therefore, the pdf of minimum over x comma y upon maximum over x comma y is equal to 2 into 1 plus z whole square 0 less than z less than 1.

Now, many of you may wonder why all the examples that I am giving are from uniform distribution or from exponential distribution why?

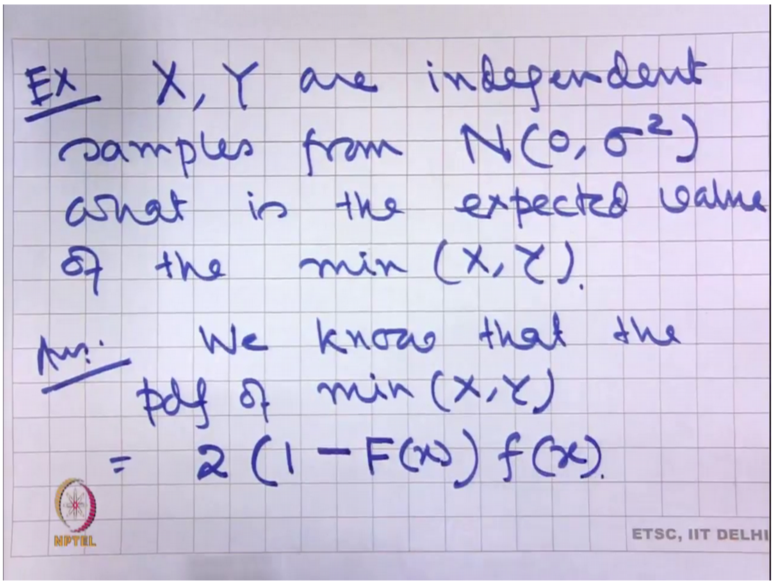
(Refer Slide Time: 25:05)



The advantage of uniform and exponential distributions or both of them have closed form for there cdf and we know that in the distribution of order statistics cdf or capital fx plays very important role because it comes in the pdf of the (Refer time: 26:05) order statistic. And therefore, the advantage of these two distributions is that we can easily handle them under order statistics.

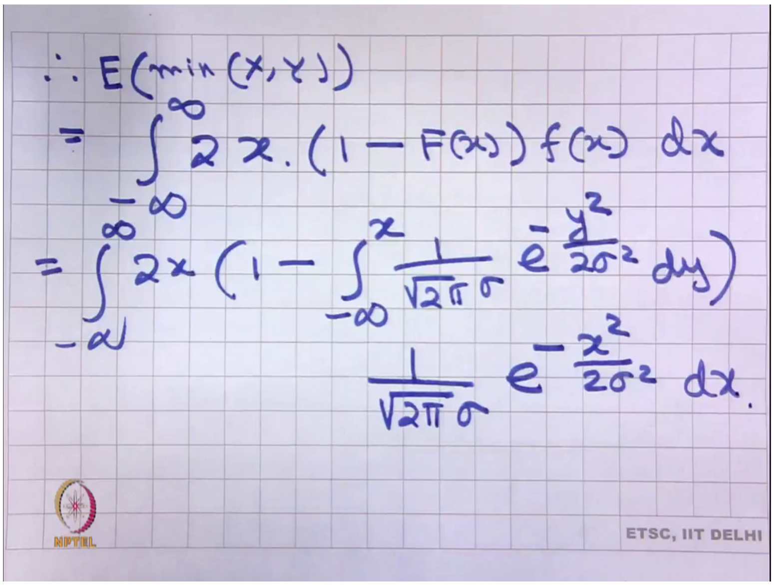
So, no wonder that most of our examples for this topic have come from uniform and exponential distribution. Now I give you a problem where I will go beyond exponential and uniform in particular I look at normal distribution and we shall see how complicated it becomes.

(Refer Slide Time: 26:50)



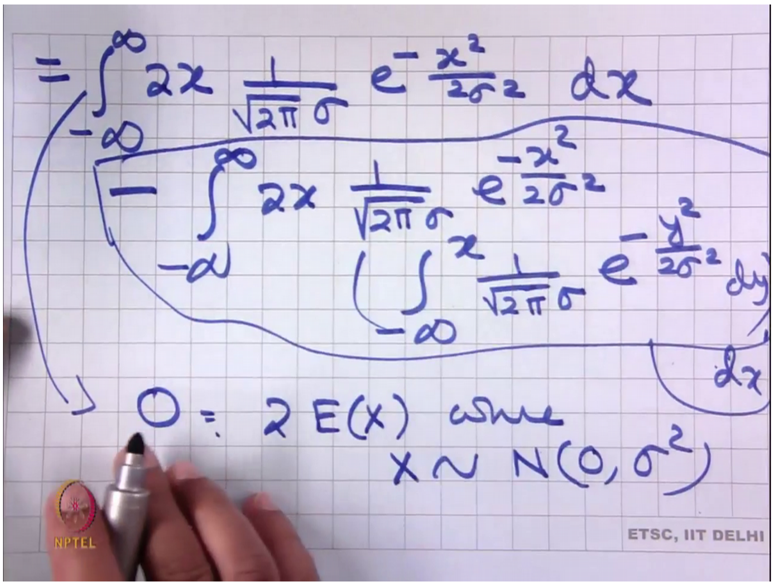
X and Y are independent samples from normal 0 sigma square. What is the expected value of the minimum of X and Y? So, that is the question. 2 into 1 minus F x into small f x.

(Refer Slide Time: 28:03)



Therefore expected value of minimum of xy is equal to integration minus infinity to infinity 2 x 1 minus F x into small f x d x is equal to minus infinity to infinity 2 x to 1 minus infinity to x 1 over root over 2 pi sigma e to the power minus y square upon 2 sigma square d y multiplied by 1 over root over 2 pi sigma e to the power minus x square upon 2 sigma square d x.

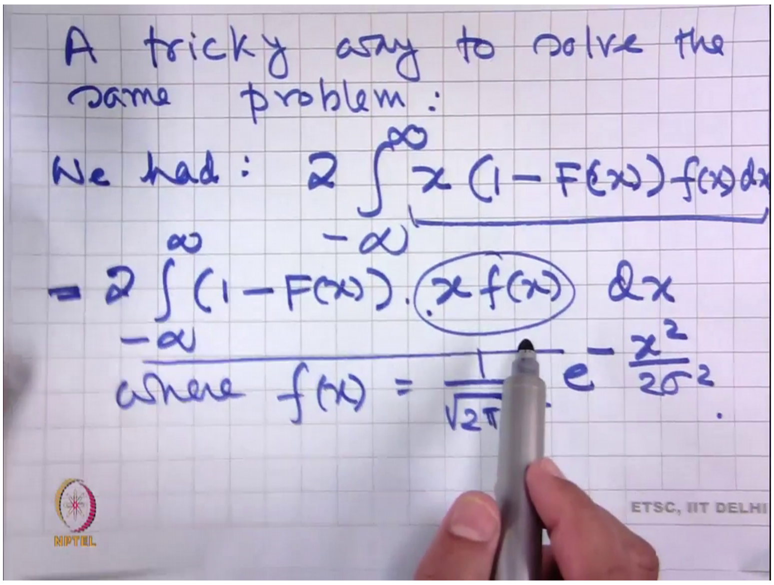
(Refer Slide Time: 29:25)



So, this I write as a sum of two parts 2 x into 1 over root over 2 pi sigma e to the power minus x square upon 2 sigma square dx I have taken out the 1 and this gives this minus minus infinity to infinity 2 x into 1 over root over 2 pi sigma e to the power minus x square 2 sigma square. This I have written first multiplied by the integration of minus infinity to x 1 over root over 2 pi sigma e to the power minus y square upon 2 sigma square dy and then whole thing is dx.

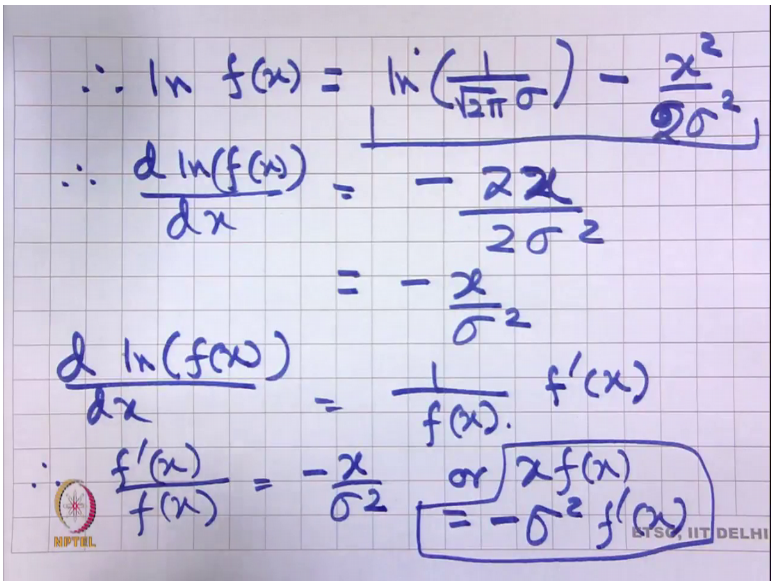
So, now you understand how complicated this expression becomes. Although this part is very very simple because this part gives us 0, why? Because this is 2 times expectation of x where x is normal with 0 sigma square. Therefore, whole concentration goes now to this term and that is complicated I argue to work up on this one and see your mathematics is strong enough to handle this thing. But what I will show you today are shortcut or a tricky way to solve the same problem.

(Refer Slide Time: 31:35)



So, we were at this stage 2 times minus infinity to infinity x into 1 minus Fx into fx dx. Which we write as follows 2 times minus infinity to infinity 1 minus Fx into x times fx dx where f is 1 over root over 2 pi sigma e to the power minus x squared upon 2 sigma square.

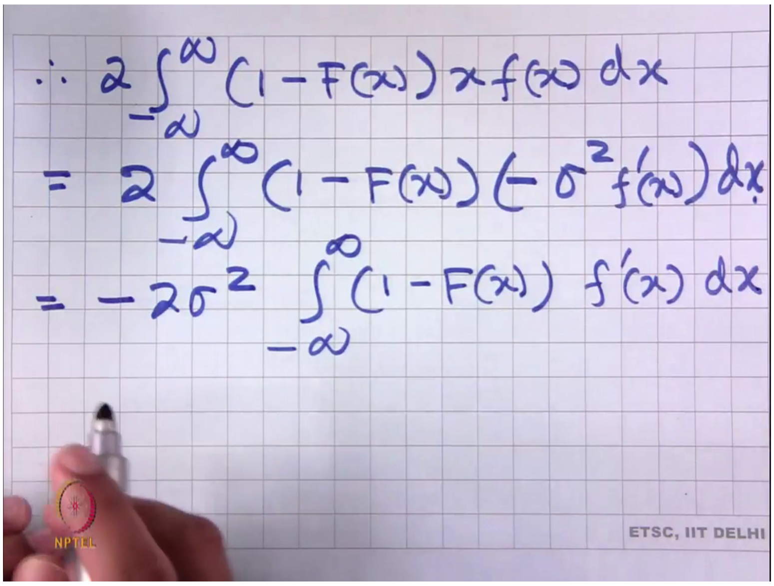
(Refer Slide Time: 32:54)



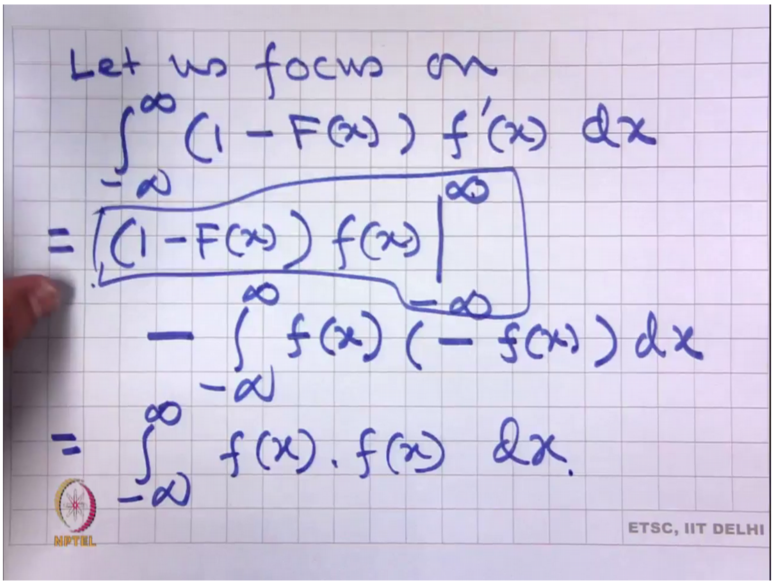
Therefore log of fx is equal to log of 1 upon root over 2 pi sigma minus x squared upon 2 sigma square therefore, d log f x dx is equal to we are differentiating this with respect to x, this becomes 0 because it is a constant and what we are getting is minus 2 x upon 2 sigma square is equal to minus x upon sigma square.

Now, on the left hand side we get d log fx dx is equal to 1 over f x into f prime x. Therefore, f prime x upon fx is equal to minus x upon sigma square or x fx is equal to minus sigma square f prime x. So, we got an expression involving x fx which we can now replace with minus sigma square f prime x. So, let us go back to our expression, we had 2 into this which involves x time fx and we know that we can replace it with minus sigma square into f prime x.

(Refer Slide Time: 35:14)

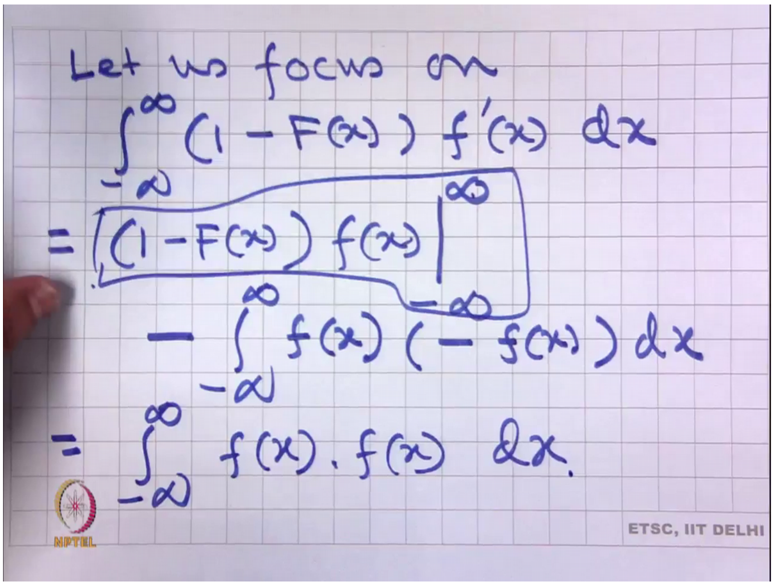


(Refer Slide Time: 36:05)



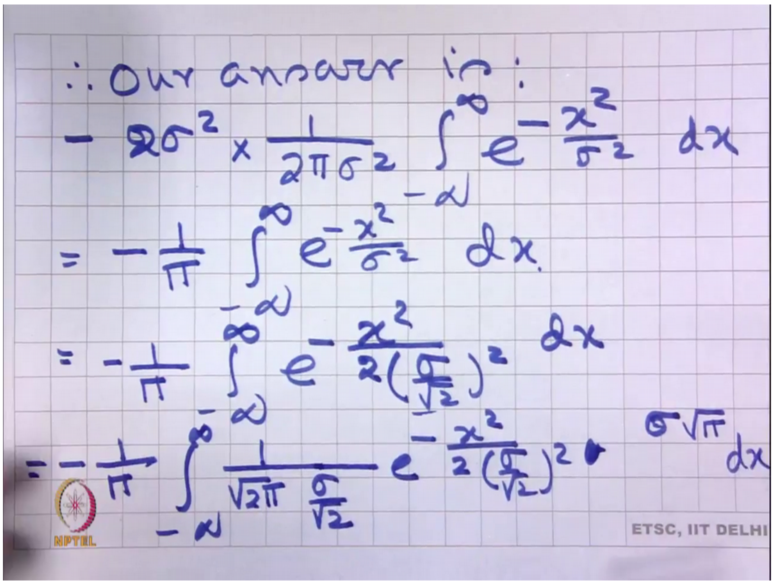
So, let us focus on minus infinity to infinity 1 minus F x into a prime x dx. This is equal to first function into integration of second function and these we put the limit minus infinity to infinity. Minus minus infinity to infinity fx now derivative of the first function is minus fx dx is equal to now the first part that x is equal to infinity capital Fx is equal to 1 therefore, it is 0 and at x is equal to minus infinity all small fx is equal to 0 as the normal distribution tappers on both sides therefore, the first part is 0. Therefore, what we are getting is integration minus infinity to infinity f x into fx dx.

(Refer Slide Time: 37:52)



Is equal to integration minus infinity to infinity 1 over root over 2 pi sigma e to the power minus x square, to sigma square multiplied by 1 over root over 2 pi sigma dx is equal to 1 over 2 pi sigma square minus infinity to infinity e to the power minus 2 x square 2 sigma square dx is equal to 1 over 2 pi sigma square into minus infinity to infinity e to the power minus x square upon sigma square d x. And remember that we started with minus to sigma square into minus infinity to infinity 1 minus fx into a prime x dx.

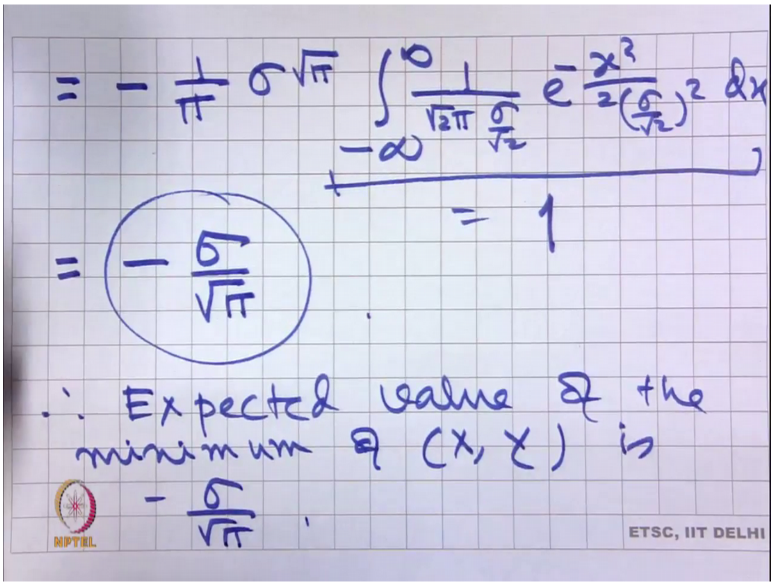
(Refer Slide Time: 39:21)



Therefore our answer is minus 2 sigma square into 1 over 2 pi sigma square into minus infinity to infinity, e to the power minus x square upon sigma square dx is equal to minus 1 upon pi into minus infinity to infinity e to the power minus x square upon sigma square dx is equal to minus 1 upon pi minus infinity to infinity e to the power minus x square upon 2 into sigma over root 2 whole square dx. Why I did it? I wanted to bring it into 2 sigma square form and therefore, I had multiplied with 2 and divided it by 2 and when I put it inside the square it becomes sigma over root 2.

What is this? If we multiply it with root over 1 over root over 2 pi sigma over root 2, then that is going to give me the density function of a normal distribution. So, let me write it as 1 over pi minus infinity to infinity 1 over root over 2 pi sigma over root 2 e to the power minus x square into 2 sigma over root 2 whole square multiplied by to compensate for this i write sigma root pi because these cancels and this gives me sigma root pi dx.

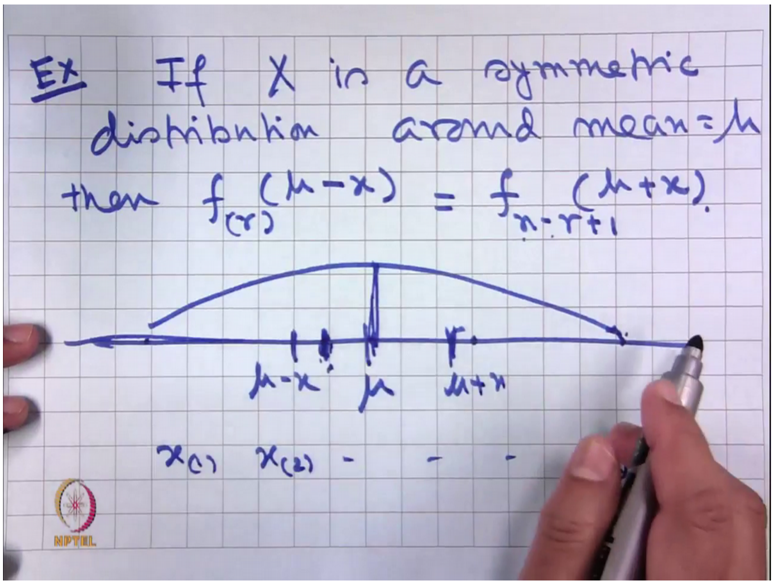
(Refer Slide Time: 41:57)



Is equal to minus 1 over pi sigma root pi integration minus infinity to infinity 1 over root over 2 pi sigma over root 2 into e to the power minus x square upon 2 sigma over root 2 whole square dx and this part is equal to 1 because I am integrating a pdf from minus infinity to infinity therefore, this is equal to minus sigma over root pi. So, that is the answer.

Now, you can see that since we were dealing with a normal distribution, things have not been as straightforward as we had with respective uniform or exponential distribution. But we have solved the problem in a tricky way and there is a longer way of doing it which I have done half way and left it as a exercise for you to solve it at home.

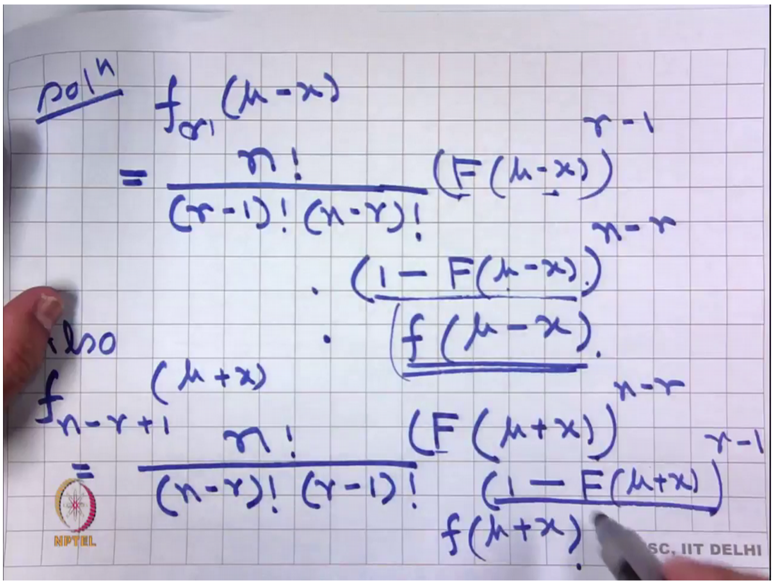
(Refer Slide Time: 43:35)



Now let me give you another tricky problem show that if X is a symmetric distribution around mean is equal to mu, then f r at mu minus x is equal to f of n minus r plus 1 at mu plus x that is x is a distribution which is symmetric around the mean mu. So, let us assume some arbitrary just pdf and suppose we have taken n samples x 1 let me order them x 2 x n.

So, if the pdf of the rth order statistics from the beginning at mu minus x will be same as the n minus r plus 1 at statistics or the rth order statistics from this side at mu plus x if sorry if this is mu minus x this is mu plus x. These two will have the same density function. How do you prove it mathematically?

(Refer Slide Time: 45:29)

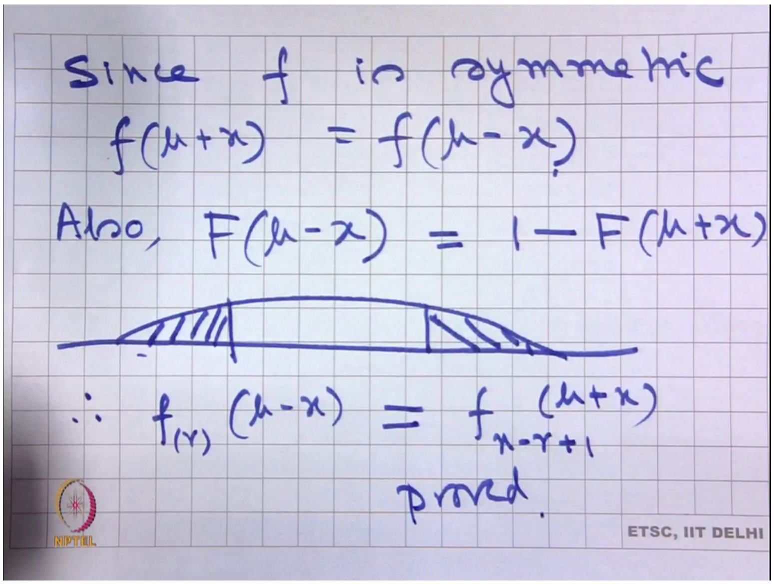


Now f r at mu minus x is equal to n factorial upon r minus 1 factorial into n minus r factorial F at nu minus x whole to the power r minus 1, 1 minus F nu minus x at n minus r multiplied by f at nu minus x. This we know from our earlier classes that first a r minus 1 has to be less than nu minus x and that probabilities F mu minus x multiplied by r minus 1 the first r minus 1 have been taken care of the rth1 at mu minus x. That we know that will be given by f mu minus x and the remaining n minus r all of them should be above mu minus x. So, why my 1 minus f mu minus x whole to the power n minus r.

Also f n minus r plus 1 at mu plus x is equal to n factorial. We are imitating this instead of what now I am putting n minus r plus 1. So, this is going to be n minus r factorial instead of or I am putting n minus r plus 1 here. So, that is going to be r minus 1 factorial f at mu plus x whole to the power n minus r because n minus r plus 1. So, one will be less so it is n minus r 1 minus f at mu plus x how many will be on this side there will be only r minus 1 of them because I am looking at the rth from the end multiplied by f at nu plus x.

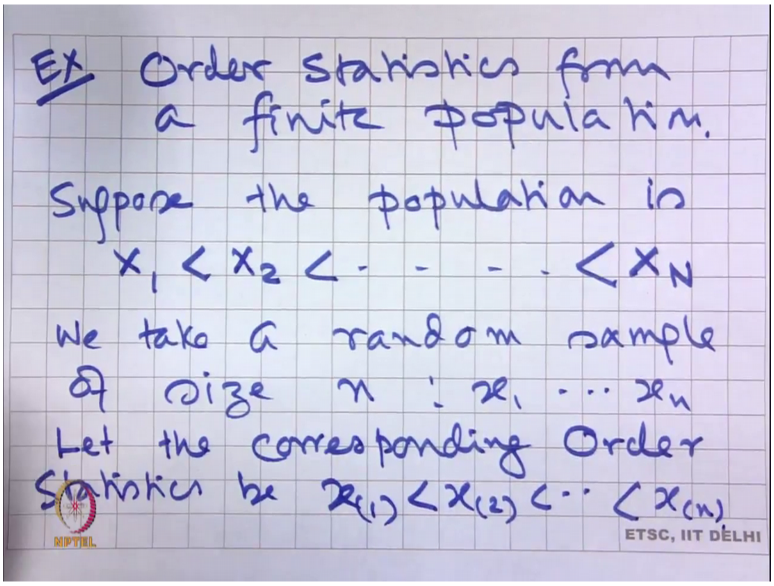
So, we got the two expressions here.

(Refer Slide Time: 48:34)



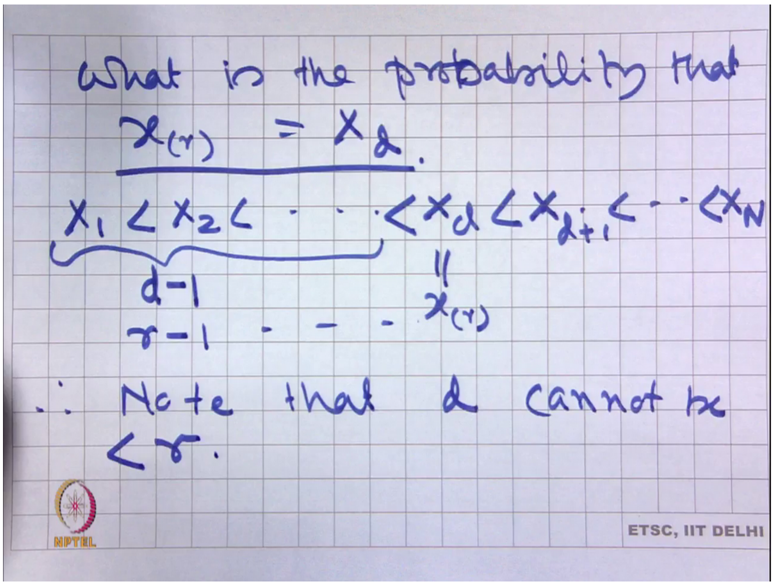
Since f is symmetric around mu f of mu plus x is equal to f of mu minus x. Also F at mu minus x is equal to 1 minus f of mu plus x. That is very simple because it is symmetric F of mu minus x is this and F of mu plus x is this. Therefore, 1 minus therefore, F of mu minus x is equal to 1 minus F of mu plus x therefore, if I put instead of F of mu minus x 1 minus F of mu plus x and instead of F of mu plus x 1 minus F nu minus x. We can see that these two terms are same that is therefore, f r at mu minus x is equal to f of n minus r plus 1 at nu plus x.

(Refer Slide Time: 50:25)



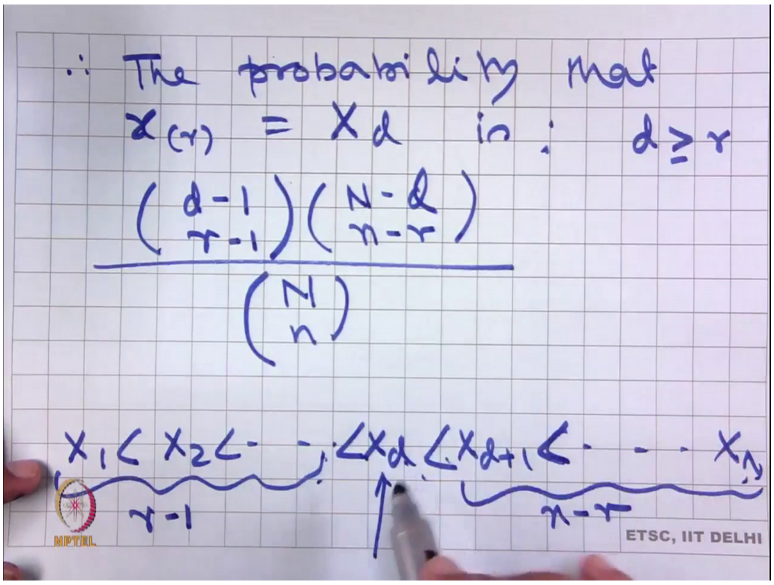
I will stop with one more problem and this time will be looking at Order Statistics from a finite population. Suppose the population is x 1; that means, we have a population of n items and they put them in increasing order X 1 X 2 xN. We take a random sample of size n x 1 x 2 xn let the corresponding order statistic be x 1 x 2 x n. And then we want to compute some of the probabilities.

(Refer Slide Time: 51:59)



So, what we are looking at, that in the sample the rth order statistic is actually the dth element in the sequence. Since this part is sorted and I am looking at that the rth order statistic is same as this what is going to be the probability. Since rth order statistics is equal to xd. We have exactly d minus 1 element on this side and in the sample, we have r minus 1 element below this. Therefore, note that d cannot be less than r or in other words suppose I am looking at ten samples from 1 2 3 up to 100. Then the fifth order statistic cannot be less than 5 it has to be at least 5 right. So, that is what I am saying here that the rth order statistic is equal to xd means they cannot be less than r.

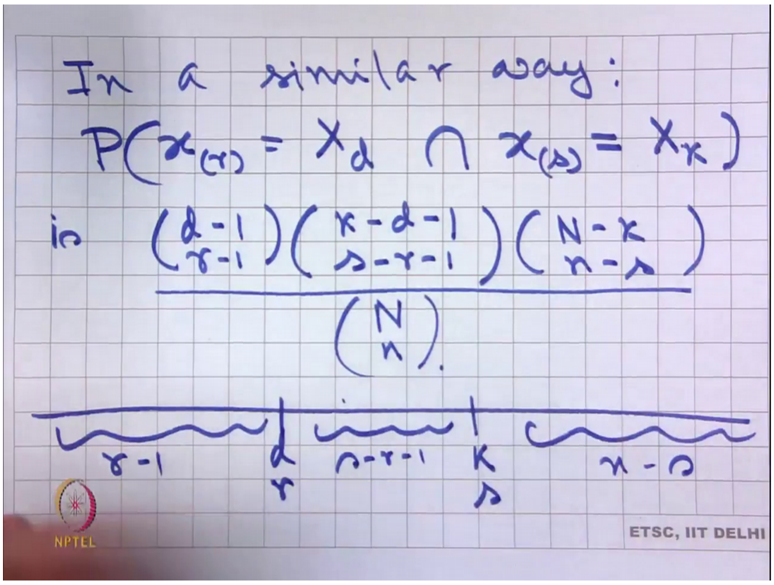
(Refer Slide Time: 54:03)



So, what is the probability? So, among the first d minus 1 I have to choose r minus 1 in d minus 1 c r minus 1 ways right. So, x 1 less than x 2. So, from here I have to choose r minus 1 and from here I have to choose n minus r and 1 will get from xd. So, this can be chosen in d minus 1 c r minus 1 ways, Here how many elements are there N minus d elements and from here I have to choose n minus r elements and the number of possible ways of selection is N cn.

Therefore this is going to be the probability that the rth order statistics of the sample is actually get one of the population.

(Refer Slide Time: 55:50).



In a similar way probability x or is equal to X d and x s is equal to X k is d minus 1 c r minus 1 multiplied by k minus d minus 1 c s minus r minus 1 multiplied by N minus k c n minus s upon N c n.

This is very simple a very similar logic will tell you that if this is the d and this is the kth 1. Then I need r minus 1 of them to be here the rth one to be here, the sth one to be here n minus s once to be here and s minus r minus 1 of them to be from here.

So, the number of ways a possible selection is coming out to be this. I want you to verify this one and I want you to attempt the problems that will be giving you in the tutorial sheet. With that I close the chapter on Order Statistics, from the next class I will go into properties of estimators.

Thank you so, much.