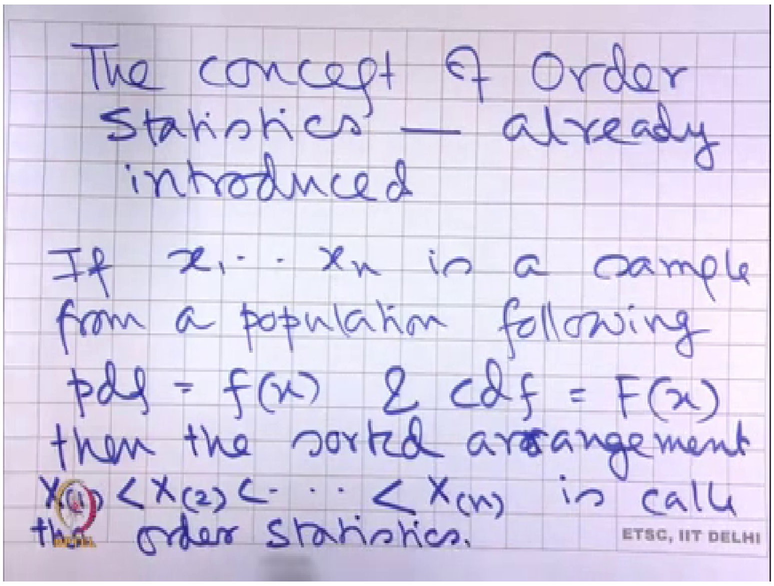
Statistical Inference

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

Lecture – 10
Statistical Inference

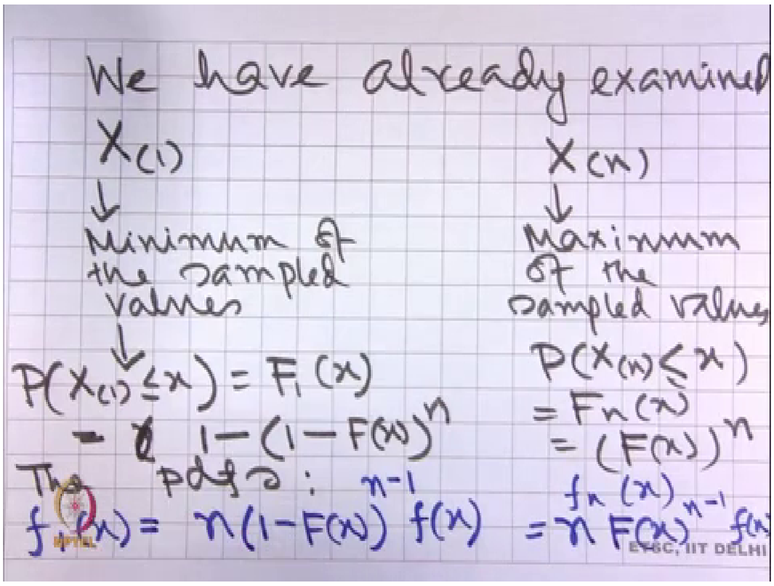
Welcome students to the MOOCS series of lectures on statistical inference, this is lecture number 10. In the last class, we have introduced the concept of order statistics.

(Refer Slide Time: 00:37)



What is that? If x 1 x 2 x n is a sample from a population following pdf is equal to f x and cdf is equal to F x then this sorted arrangement X 1 less than X 2 less than X n is called the order statistics.

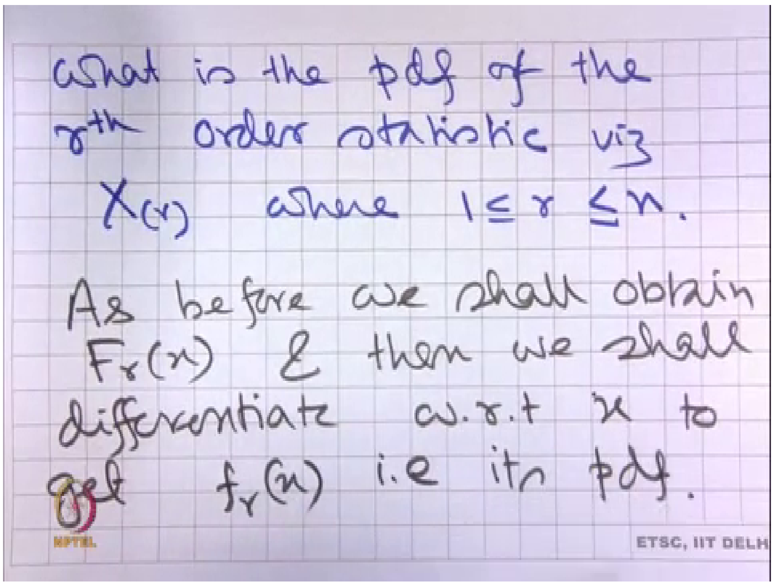
(Refer Slide Time: 02:07)



In fact, we have already examined X 1 which is the minimum of the sampled values and also X n which is maximum of the sampled values. We have also seen that the cdf probability X 1 less than equal to x is equal to F 1 x is equal to 1 minus 1 minus F x whole to the power n.

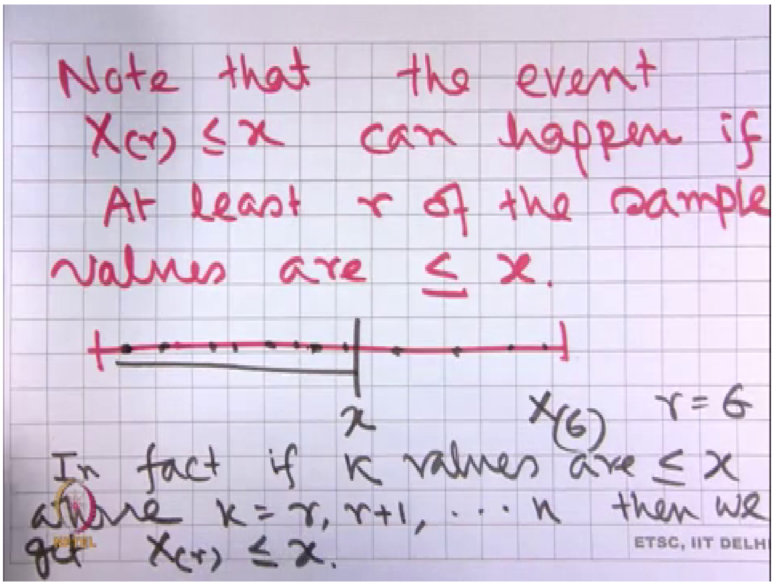
And in this case probability x nn, probability X n less than equal to x is equal to F n x is equal to F x to the power n. And from there we found the pdfs f 1 x is equal to n into 1 minus F x whole to the power n minus 1 into small f x. And in this case what we had f n x is equal to n into F x to the power n minus 1 into f x.

(Refer Slide Time: 04:39)



This we have already seen and today I will look at what is the pdf of the r th order statistic namely X r, where 1 less than equal to r less than equal to n. As before, we shall obtain Fr x and then we shall differentiate with respect to x to fame to all to get f r x that is its pdf,.

(Refer Slide Time: 06:01)

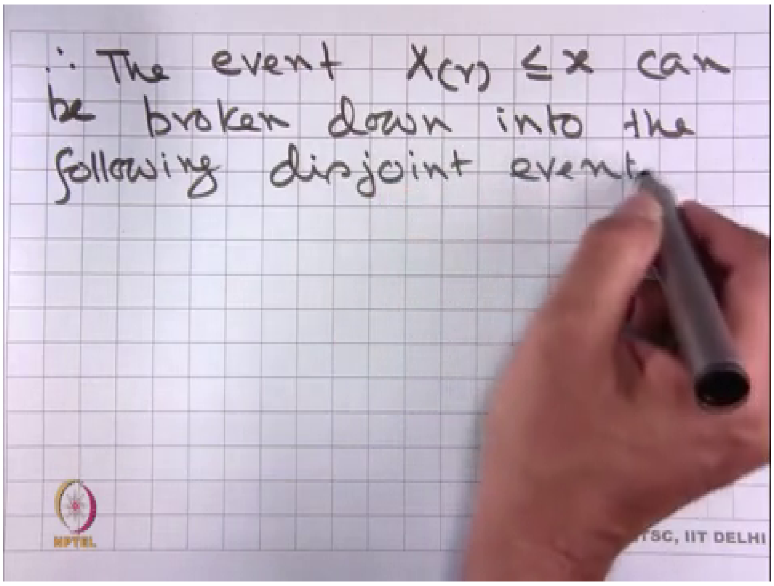


Note that the event X r less than equal to x can happen if at least r of the samples or r of the sample values are less than equal to x.

So, if these are the samples and this is x and suppose I want to find out X 6, the distribution of X 6 that is r is equal to 6. That can happen if there are 6 values less than equal to x, but that will also happen if there are 7 values less than equal to x because even in that case also the sixth one is less than equal to x. In fact, if it is 8 still X 6 will remain less than equal to 6.

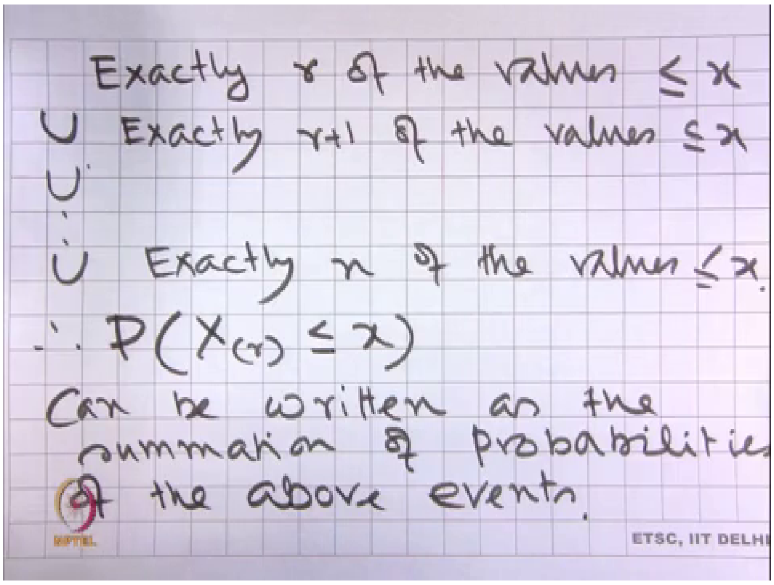
In fact, if K values are less than equal to x where k is equal to r, r plus 1 up to n then we will get X r less than equal to x.

(Refer Slide Time: 08:21)



Once we understand that X r less than equal to x can be broken down into the following disjoint events.

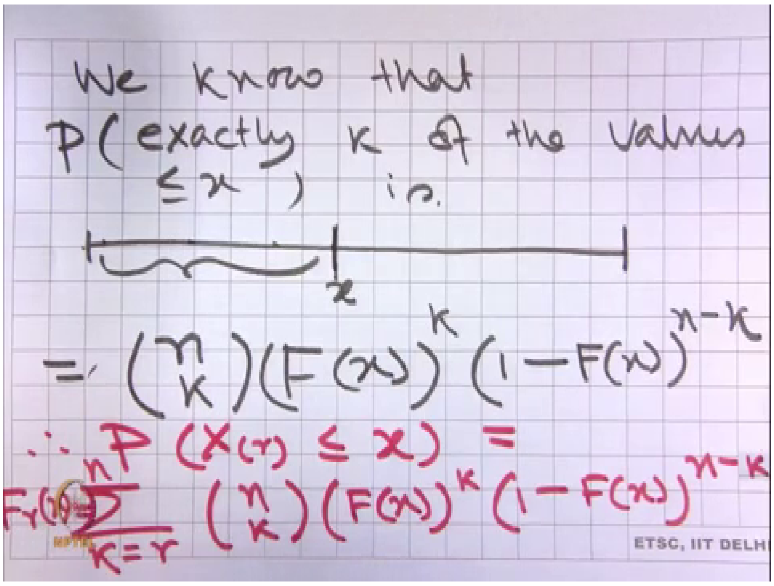
(Refer Slide Time: 08:57)



Exactly r of the values less than equal to x, Union Exactly r plus 1 of the values less than equal to x, Union Exactly n of the values less than equal to x.

Therefore probability X r less than equal to x can be written as the summation of probabilities of the above events because these are all disjoint.

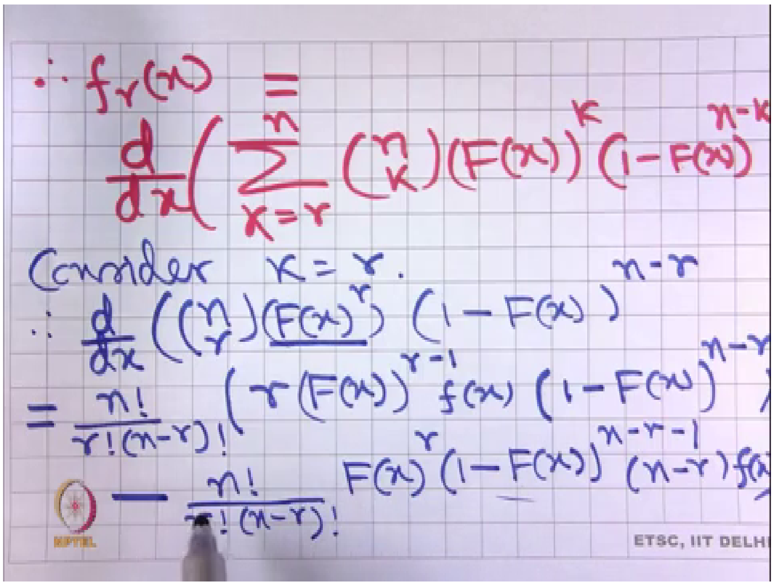
(Refer Slide Time: 10:21)



We know that Probability exactly K of the values less than equal to x is out of the n ,samples, we will choose K of them and they will be put in the interval less than equal to x. And the remaining n minus k will remain in this part which is greater than x.

Therefore, this probabilities, number of ways of choosing K out of n nck, all of them are having values less than x. So, F x to the power k and all the remaining one of them are above x. So, it is 1 minus F x whole to the power n minus k. Therefore, probability X r less than equal to x can be written as the summation K is equal to r to n nck F x to the power k 1 minus F x whole to the power n minus k. Therefore, what is going to be the pdf, so, this is F r x therefore, what is going to be the pdf, the pdf is the derivative of this with respect to x.

(Refer Slide Time: 12:52)

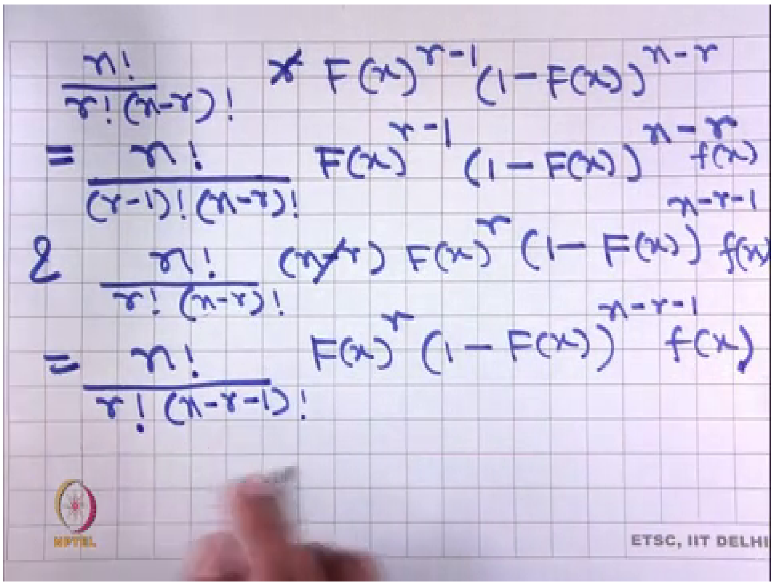


Therefore f r x is equal to d dx of sigma K is equal to r 2 n nck F x to the power k 1 minus F x whole to the power n minus k. Since this is a summation we can differentiate them term by term. So, consider K is equal to r. Therefore, we are looking at d dx of n c r F x to the power r and 1 minus F x whole to the power n minus r is equal to it is a product of two terms.

So, we have to look at derivative of first function multiplied by the second function plus derivative of the second function multiplied by the first function. Therefore, this is going to be factorial n upon factorial r into factorial n minus r that we are getting from this multiplied by now I am differentiating this with respect to x and we are getting r F x to the power r minus 1 multiplied by f x whole multiplied by 1 minus F x whole to the power n minus r.

Now, the second term is its derivative multiplied by this. Therefore, that is coming out to be factorial n upon factorial r into factorial n minus r into F x to the power r into 1 minus F x whole to the power n minus r minus 1 multiplied by n minus r, because that will come. And then I am differentiating 1 minus f x with respect to x that gives me f x with a minus sign. Thus, we get 2 terms 1 is the first part and other is the second part which is actually becoming a subtraction.

(Refer Slide Time: 16:17)



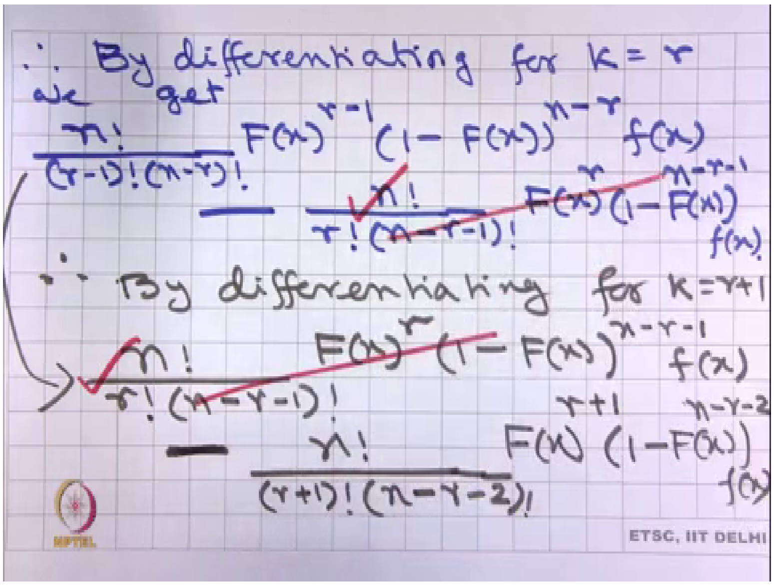
Now, let us simplify it factorial n into factorial r into factorial n minus r multiplied by r into F x to the power r minus 1 into 1 minus F x whole to the power n minus r is equal to these r cancels with this.

So, what we are getting is n factorial into r minus 1 factorial into n minus r factorial F x to the power r minus 1 into 1 minus F x whole to the power n minus r.

And what is the second component? It is n factorial into r factorial into n minus r factorial into n minus r F x to the power r 1 minus F x to the power n minus r minus 1. And of course, there will be an f x term there that should have been here also is equal to, now these cancels with the first one. So, that gives me n factorial r factorial n minus r minus 1 factorial this cancels F x to the power r 1 minus F x whole to the power n minus r minus 1 multiplied by f x.

Thus, the first term produces the following.

(Refer Slide Time: 18:33)

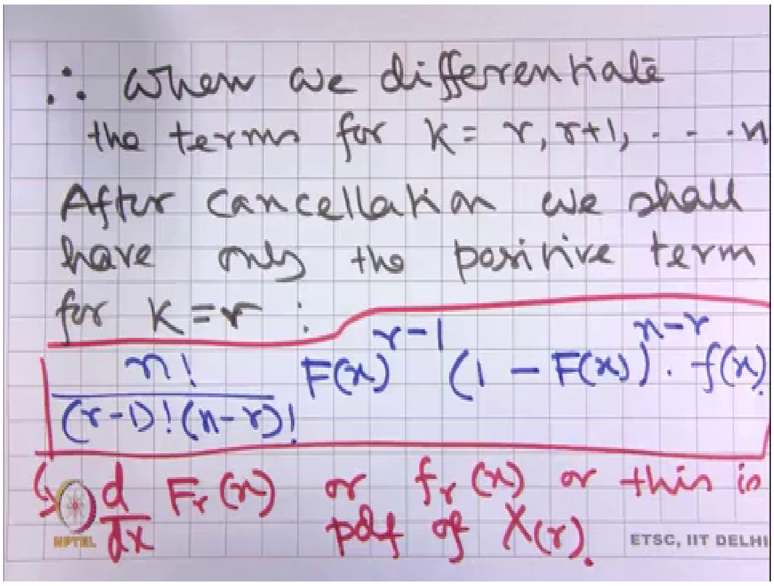


Therefore, by differentiating for K is equal to r, we get n factorial upon r minus 1 factorial into n minus r factorial into F x to the power r minus 1 into 1 minus F x to the power r n minus r into F x minus n factorial r factorial n minus r minus 1 factorial F x to the power r 1 minus F x to the power n minus r minus 1 into f x.

Therefore by differentiating for k is equal to r plus 1, what we shall get. We shall get a very similar expression where r is replaced with r plus 1. Therefore, what we are writing you just notice I will be replacing everywhere r with r plus 1. So, what we will get is n factorial, it is r factorial it is n minus r minus 1 factorial, because I am subtracting r plus 1 now F x to the power r plus 1 minus 1 which is r 1 minus F x whole to the power n minus r minus 1 into f x, subtraction n factorial r plus 1 factorial n minus r minus 2 factorial F x to the power r plus 1 r is replaced with r plus 1. And now here r will be replaced with r plus 1, so, what will get 1 minus F x whole to the power n minus r minus 2 multiplied by f x.

Now there is an interesting thing that happened, if you consider this term and this term, they are the same n factorial upon r factorial into n minus r minus 1 factorial F x to the power r 1 minus F x to the power n minus r minus 1 into small effects. So, they cancel with each other.

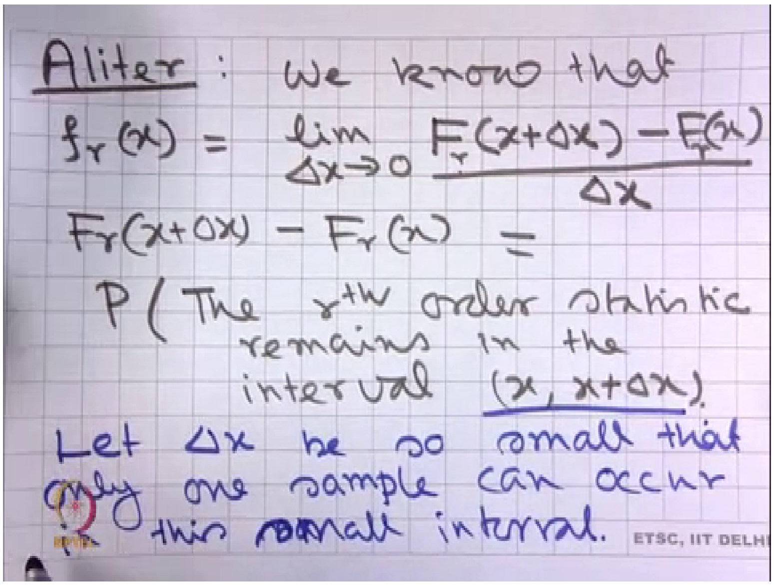
(Refer Slide Time: 22:31)



Therefore when we differentiate the terms for K is equal to r comma r plus 1 up to this; they will keep on canceling each other. So, what will remain is we shall have only the positive term for K is equal to r which is as we have just obtained factorial n upon factorial r minus 1 into n minus r factorial into F x to the power r minus 1 into 1 minus F x whole to the power n minus r into fx.

You note that for the last term n, when we differentiate their own be any negative term. Therefore, there will not be anything remaining at the end. So, this is the only thing that we get as d dx of F r x or this is f r x or this is pdf of X r. Hope you understand the calculation, but this is a bit mathematically oriented, I give you a very interesting way of looking at it in a different way.

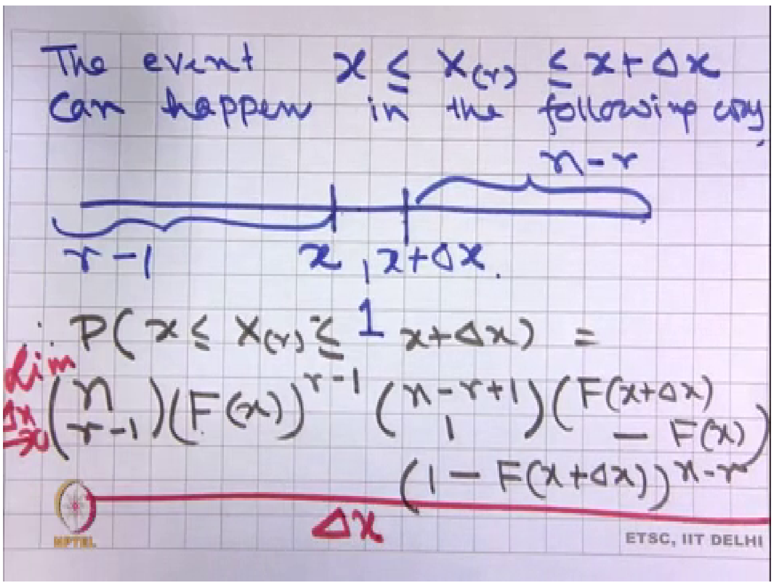
(Refer Slide Time: 25:09)



We know that f r x is equal to limit delta x going to 0 F r x plus, F r x plus delta x minus F x, F r x divided by delta x that comes from the definition of differentiation.

Now, what is F r x plus delta x minus F r x. This is probability that the rth order statistic remains in the interval x comma x plus delta x. Let delta x be so small that only 1 sample can occur in this small interval x 2 x plus delta x. Then what will happen.

(Refer Slide Time: 27:35)

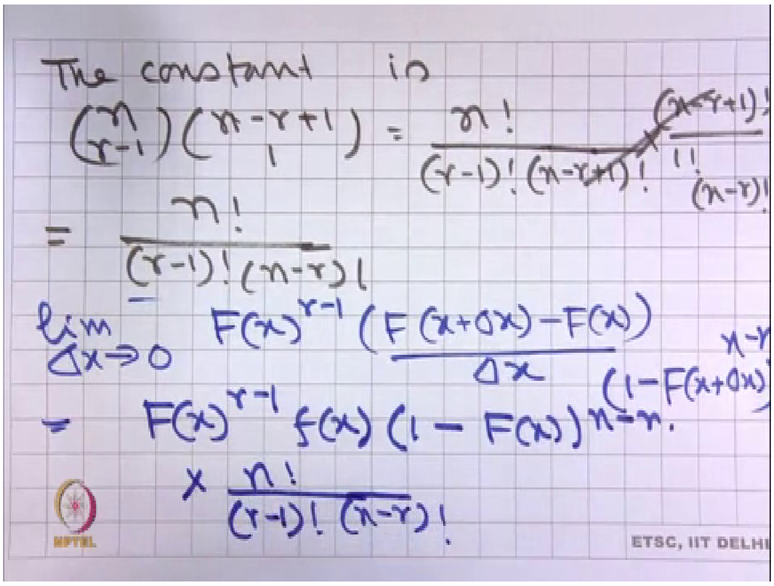


The event x less than equal to X r less than equal to x plus delta x can happen in the following way. Suppose this is my interval, this is x, this is x plus delta x.

So, we are looking at X r. So, X r will occur in this interval. So, only 1 among the n observations will be here. Since it is the rth 1 r minus 1 observations will be less than x and n minus r observations will be above x plus delta x. How can that happen, what is the probability? Out of n, I am choosing r minus 1 of them and they are less than equal to x therefore, that probabilities F x to the power r minus 1 after I have chosen r minus 1, I am left with n minus r plus 1 elements. Out of that one has been chosen to put it between x to x plus delta x and that probability will be F x plus delta x minus F x, here F x is the cdf from the parent population.

Then whatever is remain n minus r all of them will be in this interval and that probability is 1 minus f of x plus delta x whole to the power n minus r. Now we have to defer, divide it by delta x and we need to take the limit delta x going to 0 to find the derivative. So, we will be looking at limit delta x going to 0 of this divided by delta x. What is that? This is going to be the following.

(Refer Slide Time: 31:10)

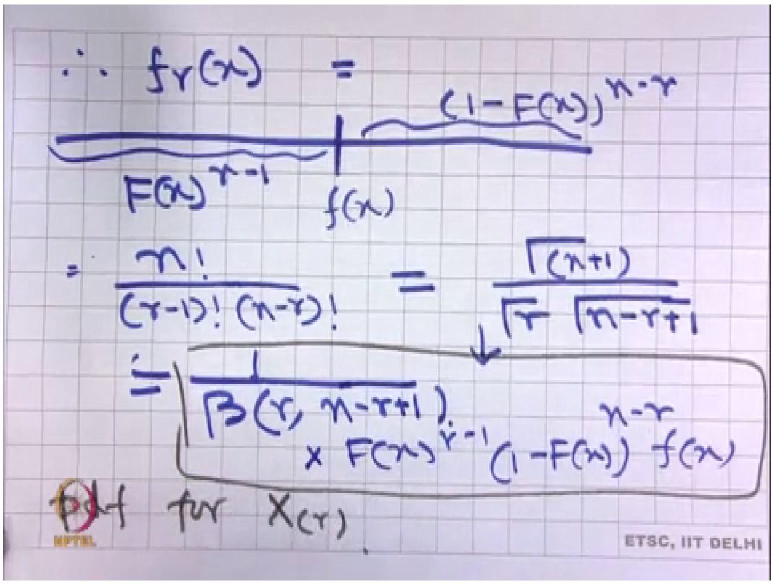


The constant is n c r minus 1 into n minus r plus 1 c 1 is equal to factorial n factorial r minus 1 into factorial n minus r plus 1 into factorial n minus r plus 1 factorial 1 into factorial n minus r is equal to n factorials upon r minus 1 factorial into n minus r factorial.

So, this is the constant term. Now let us look at limit delta x going to 0 of F x to the power r minus 1 F of x plus delta x minus F x upon delta x and multiplied by 1 minus F of x plus delta x whole to the power n minus r.

So, if we take the limit delta x going to 0 then what we are getting F x to the power r minus 1, this is going to give us f x and as delta s goes to 0 this becomes 1 minus F x whole to the power n minus r and this whole thing will be multiplied by the constant that we have obtained here that is n factorial into r minus 1 factorial into n minus r factorial. Therefore, how do you remember this? It is very simple.

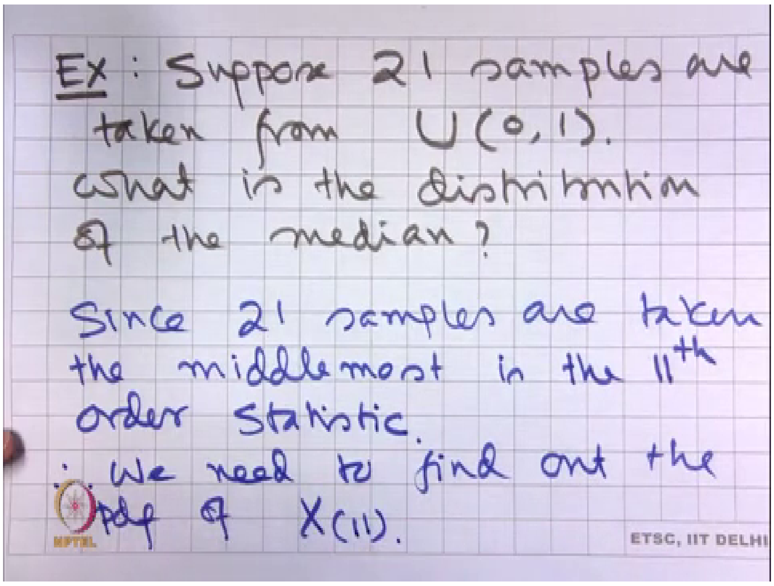
(Refer Slide Time: 33:23)



So, f r x is equal to if this is x so, we are putting one here with Fx in this part there are r minus 1 that is giving me f x to the power r minus 1 in this part there are n minus r. So, that is giving me 1 minus F x whole to the power n minus r. And the whole thing will be multiplied by the constant which is factorial n upon factorial r minus 1 into factorial n minus r.

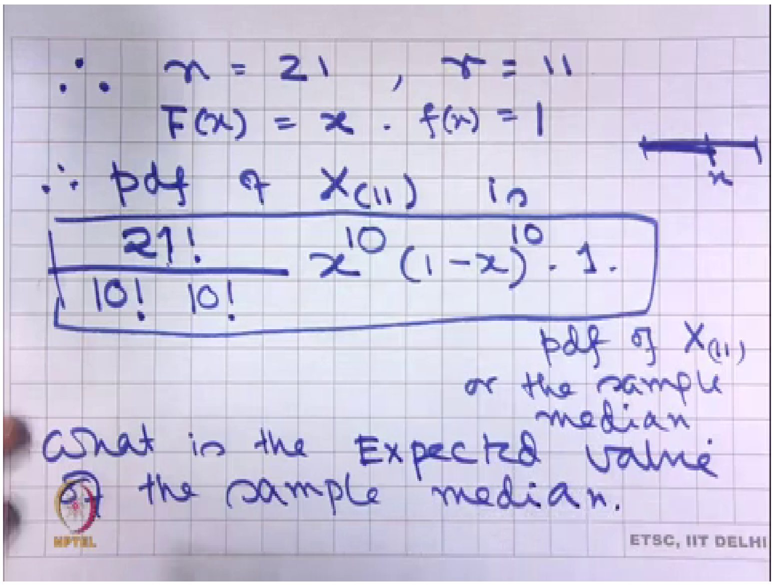
This is not exactly the binomial coefficient, but we know that gamma n plus one is equal to factorial n. Therefore, if we write it as Gamma n plus 1, this is Gamma r and this is Gamma n minus r plus 1 then we see that this comes under standard Beta. In fact, it is 1 upon Beta r comma n minus r plus 1 therefore, 1 upon Beta or Gamma n minus r plus 1 into F x to the power r minus 1 into 1 minus F x to the power n minus r into fx, this is the pdf.

.(Refer Slide Time: 35:45)



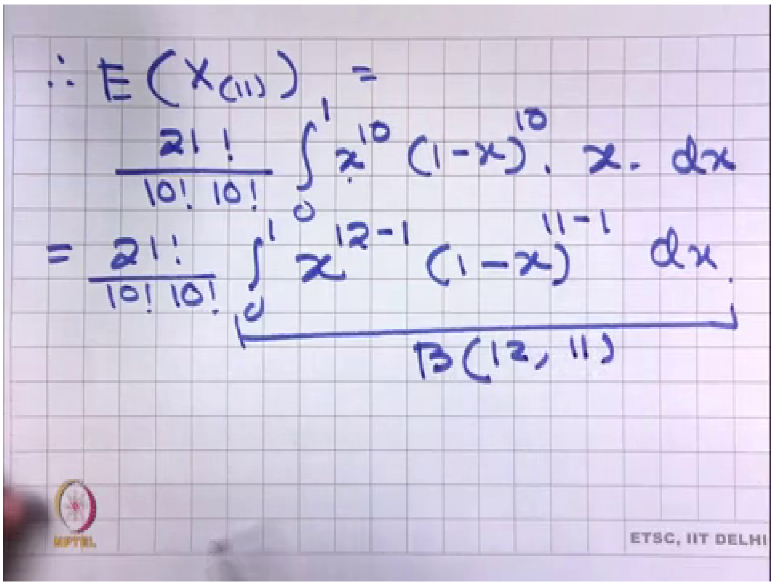
Let me solve a problem. Suppose 21 samples are taken from uniform 0, 1. What is the distribution of the median? We know that since 21 samples are taken, the middlemost is the 11 th order statistic Therefore, we need to find out the pdf of X 11.

(Refer Slide Time: 37:27)



Therefore n is equal to 21, r is equal to 11, F x is equal to x because it is uniform 0, 1. So, probability less than equal to x is equal to x and f x is equal to 1. Therefore, pdf of X 11 is factorial 21 factorial r minus 1 r is equal to 11 so, it is 10 factorial, it is 21 factorial. it is 10 factorial, factorial n minus r n is 21 r is 11 therefore, it is 10 factorial then fx to the power r minus 1 what is F x? F x is the capital F x is x therefore, x to the power 10 1 minus x to the power n minus x which is 10 multiplied by f x which is 1

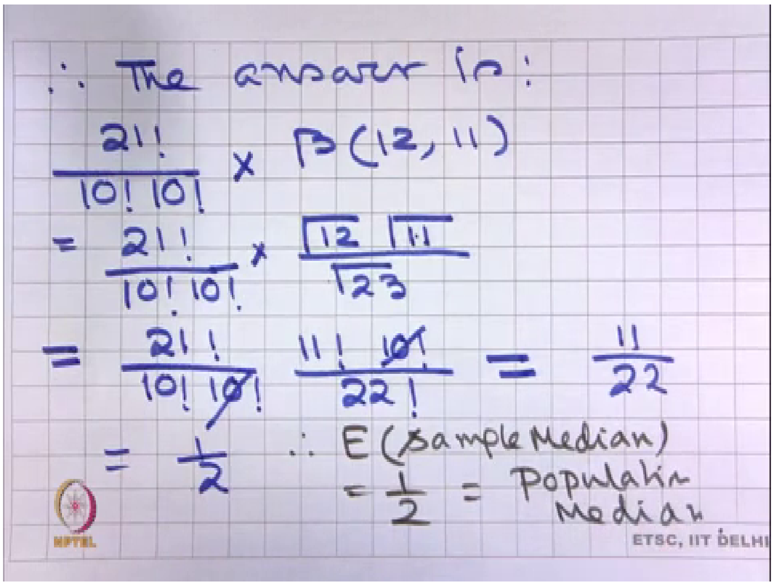
.(Refer Slide Time: 39:41)



Therefore this is the pdf of X 11 or the sample median. Therefore, if I ask you what is the expected value of, of the sample median? What will do? Expected value of X 11 is equal to 21 factorial 10 factorial 10 factorial integration 0 to 1 x to the power 10 1 minus x to the power 10 multiplied by x into dx is equal to 21 factorials 10 factorial 10 factorial 0 to 1, it is x to the power 10 into x. So, x to the power 11.

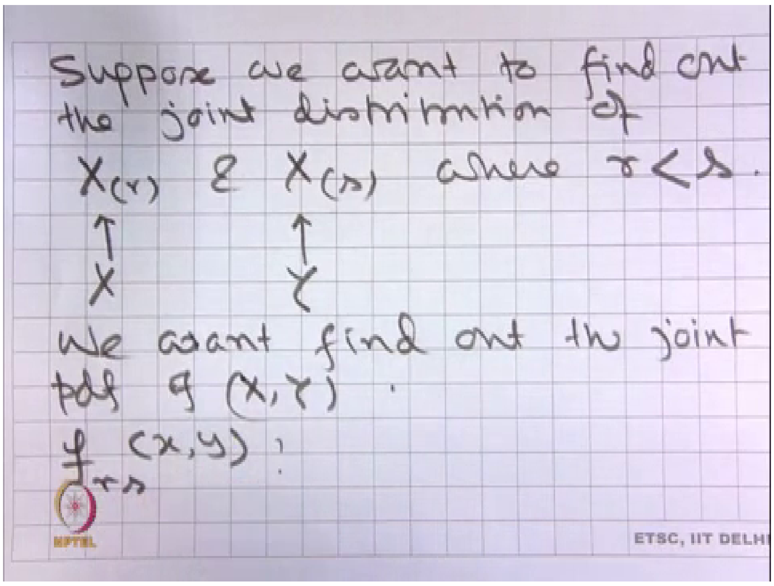
Let me write it as x to the power 12 minus 1 into 1 minus x it is 10. So, let me write it as 11 minus 1 dx, why I did it because then it becomes Beta 12 comma 11.

(Refer Slide Time: 40:57)



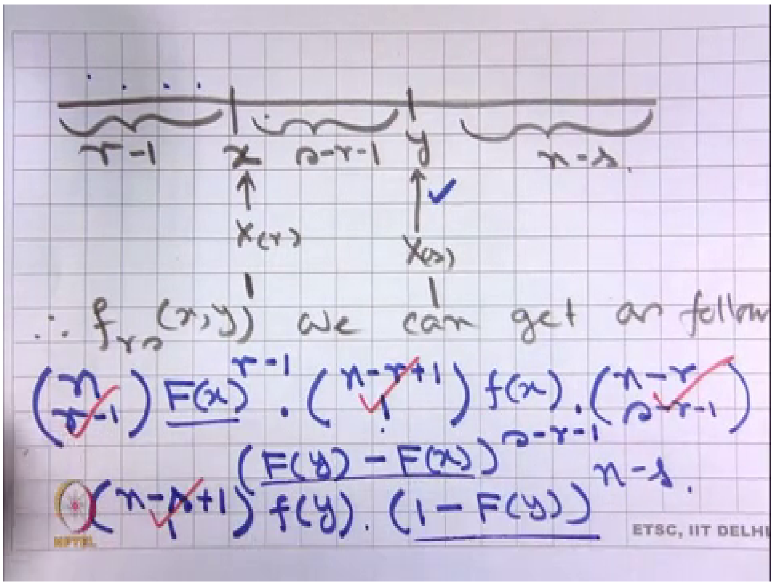
Therefore, the answer is 21 factorial 10 factorial 10 factorial into Beta 12 comma 11 is equal to 21 factorial 10 factorial 10 factorial into Gamma 12 Gamma 11 upon Gamma 23 Gamma m Gamma n comma Gamma m plus n, we know that factorial n is equal to Gamma n plus 1. So, this we can write it as 21 factorial 10 factorial 10 factorial, this is 11 factorial, this is 10 factorial, this is 22 factorial is equal to 11 upon 22 is equal to half. Thus, expected value of sample median is equal to half, is equal to Population Median.

(Refer Slide Time: 42:50)



Now, we can go slightly further. Suppose we want to find out the joint distribution of X r and X s. So, it is the rth order statistic, it is the sth order statistic where r is less than s. Let us denote this by X and this by Y. So, effectively we are looking at the joint pdf that is f order statistic r comma s, x comma y. What is that joint pdf? We can compute it mathematically, but I will show you using the type of diagram, I used in the second way of doing when I was just looking at the rth order statistic.

(Refer Slide Time: 44:29)

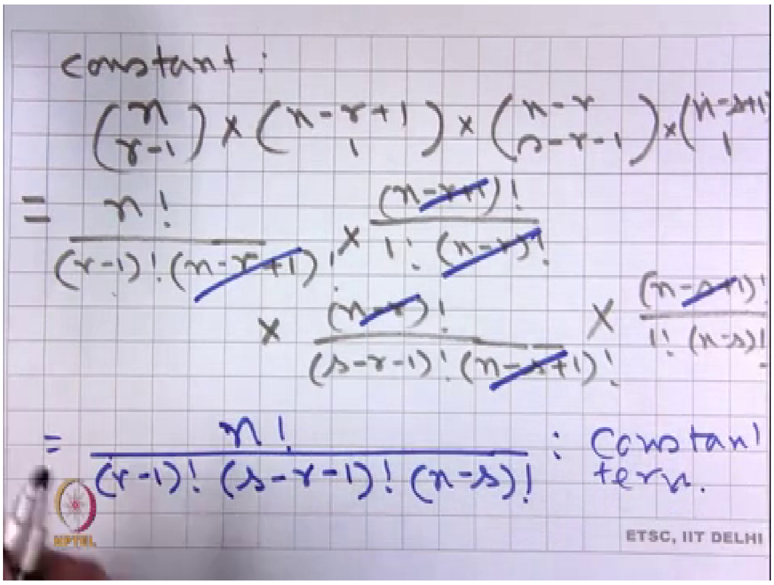


So, this is the line suppose I want the rth order of the statistic here and the sth order statistic here then what we will do how many will be here in this, there will be r minus 1 observations less than x, rth order will be at this point. How many will be here s minus r minus 1 observations will be here 1 will be here and remaining n minus s will be there.

Therefore, f r is x y, we can get as follows. Out of n I am choosing r minus 1 to fall here and that probabilities F x to the power r minus 1. Now how many r remaining, n minus r plus 1. Out of that n minus r plus 1 I am choosing 1 and putting it here, we know that for that we will get an f x we have seen it when we are doing for only rth order statistic.

Now, how many are remaining n minus r plus 1 out of that 1 I have taken. So, n minus r remaining. Out of that s minus r minus 1 are chosen to be within this interval and that probability is going to be F y minus F x whole to the power s minus r minus 1. Now how many are left n minus r minus s plus r plus 1 n minus r minus s plus r that cancels the r plus 1, from there I am choosing 1 and that I am putting in this position. So, that gives me F y. How many more are left, only n minus s are left all of them are assigned here. So, that gives me 1 minus F y whole to the power n minus s. It is apparently a complicated term. So, let me simplify it.

(Refer Slide Time: 47:57)



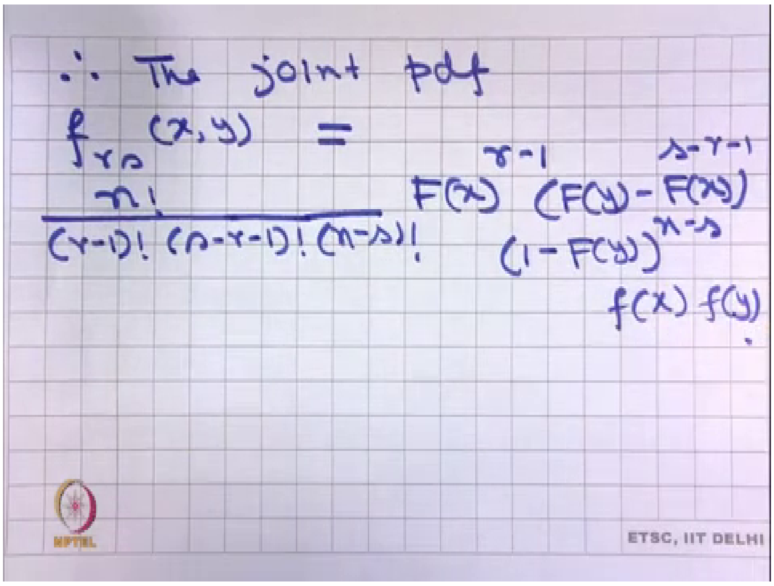
The constant is n c r minus 1 into n minus r plus 1 c 1 into n minus r c s minus r minus 1 into n minus s plus 1c 1 is equal to factorial n factorial r minus 1 factorial n minus r plus 1 multiplied by n minus r plus 1 factorial into 1 factorial into n minus r factorial into from here I am getting n minus r factorial upon s minus r minus 1 factorial into n minus s plus 1 factorial into n minus s plus 1 factorial upon 1 factorial into n minus s factorial.

It is apparently complicated, but if you understand what I am doing, it will be very clear to you. Now what cancels out, this cancels out, this cancels out and this cancels out.

So, what is remaining is n factorial upon r minus 1 factorial. It is 1 factorial so we can forget it s minus r minus 1 factorial into n minus s factorial. So, this is going to be the constant term.

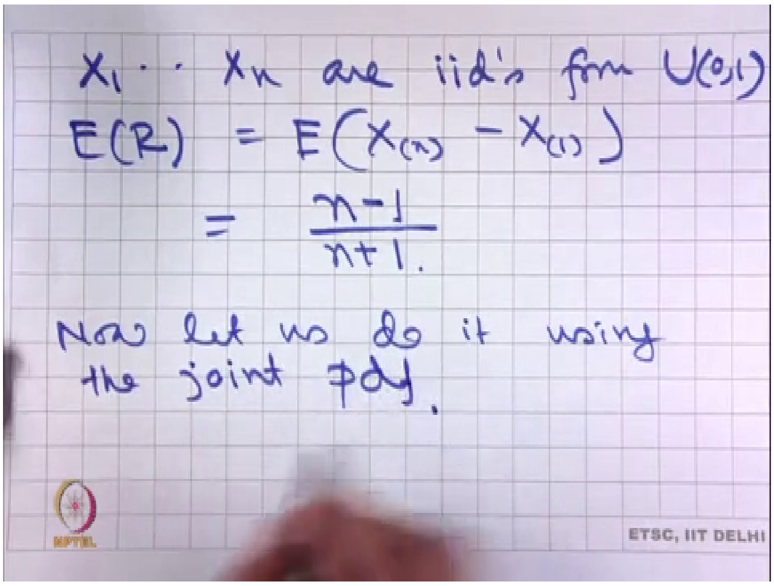
If you remember, we have done this we have used this and of which we have taken care of this, this, this, this and of course, f x and f y come together what is remaining is this and this.

(Refer Slide Time: 50:41)



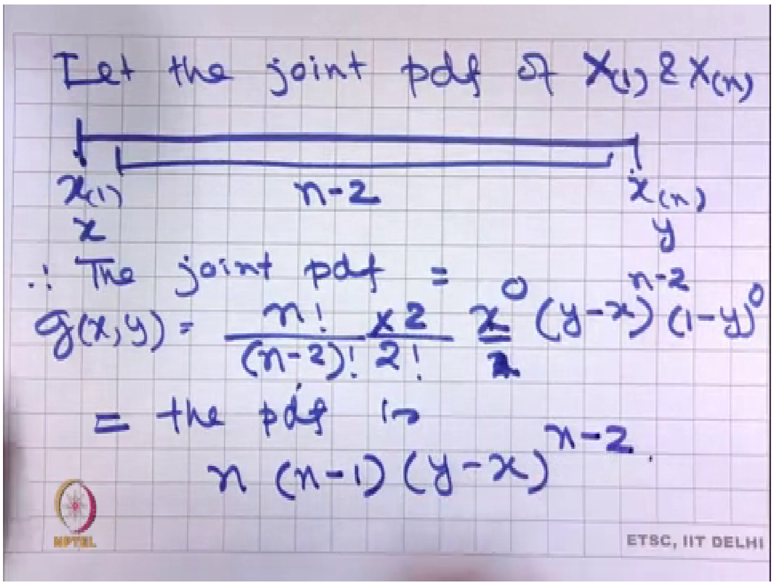
So, the joint pdf f r s x comma y is equal to n factorial r minus 1 factorial s minus r minus 1 factorial n minus s factorial into F x to the power r minus 1 into F y minus F x to the power s minus r minus 1 into 1 minus F y to the power n minus s into f x f y.

(Refer Slide Time: 52:05)



Let me give you an example. We have already seen if X 1 X 2 X n are iid is from uniform 0, 1, we calculated the expected value of range is equal to expected value of X n minus X 1 and we have found that it is, is equal to n minus 1 to n plus 1. We have not actually calculated through the pdf. Now let us look at it using the joint pdf.

(Refer Slide Time: 53:03)



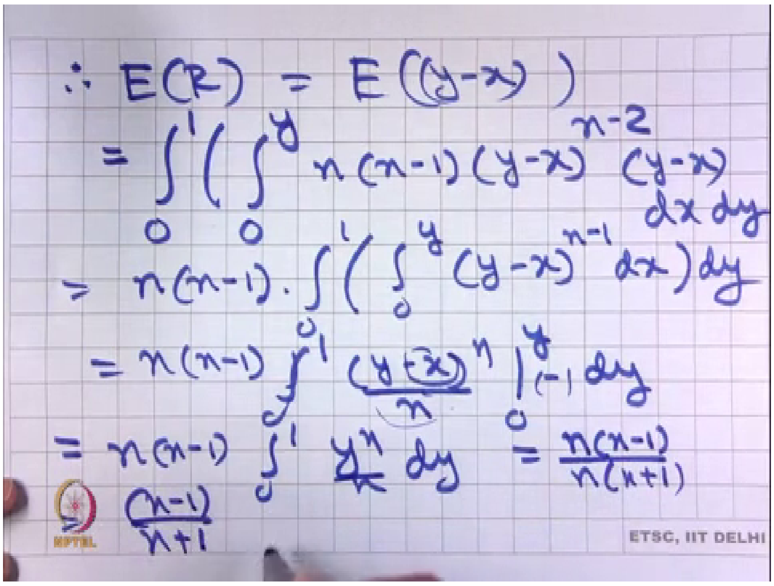
So, let the joint pdf be. So, my r is 1, my s is n I am looking at, at the 2 extremes of the 2 of the2 of the all of, we are looking at the 2 extremes of the samples.

So, my X1 is here my X n is here which I am calling at x this is I am calling y and therefore, how many are there, there are n minus 2 elements in between.

Therefore, the joint pdf is let me call it g x y is equal to n factorial into n minus 2 factorial into 2 factorial x to the power 0 because there will not be any before this x. So, x to the power 0 y minus x to the power n minus 2 and 1 minus y to the power 0 because there will not be anything here, but this I can choose in 2 possible ways.

So, that will be multiplied by a 2, out of n we have chosen n minus 2. Now from the remaining 2, I can choose in 2 possible ways 1 which will remain on this side and y 1 which will remain not at x and 1 which will remain at y.

(Refer Slide Time: 55:45)



Therefore the pdf is n into n minus 1 into y minus x whole to the power n minus 2. therefore, expected value of range is equal to expected value of y minus x is equal to we have to integrate it with respect to both x and y suppose y goes from 0 to 1 then x can go only from 0 to y because x cannot go beyond y it is the smaller of the 2 n into n minus 1 into y minus x whole to the power n minus 2 into y minus x dx dy is equal to n into n minus 1 0 to 1, 0 to y y minus x to the power n minus 1 dx dy is equal to n into n minus 1 0 to 1 y minus x whole to the power n upon n 0 to y dy is equal to n into n minus 1 integration 0 to 1.

Now, if I put y is equal to 0 then this becomes if I put x is equal to y this becomes 0 if I put x is equal to 0 it becomes y, but since it is minus x there should have been a negative sign and then that will make it positive. So, what will be remaining is y to the power n upon n dy is equal to n into n minus 1 into upon n into n plus 1 is equal to n minus 1 upon n plus 1.

So, that is the answer that we have got when we have used the linearity of expectation and we simply calculated expectation of X n minus expectation of X 1 ok. I will stop here today in the next class I shall solve a few more problems in order statistic. So, that you understand the concept and you will be able to solve the problems given in the tutorials.

Thank you.