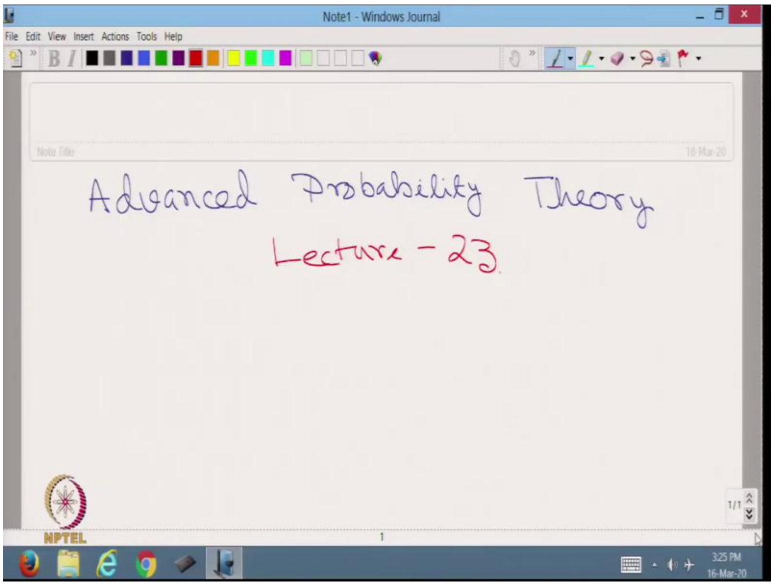
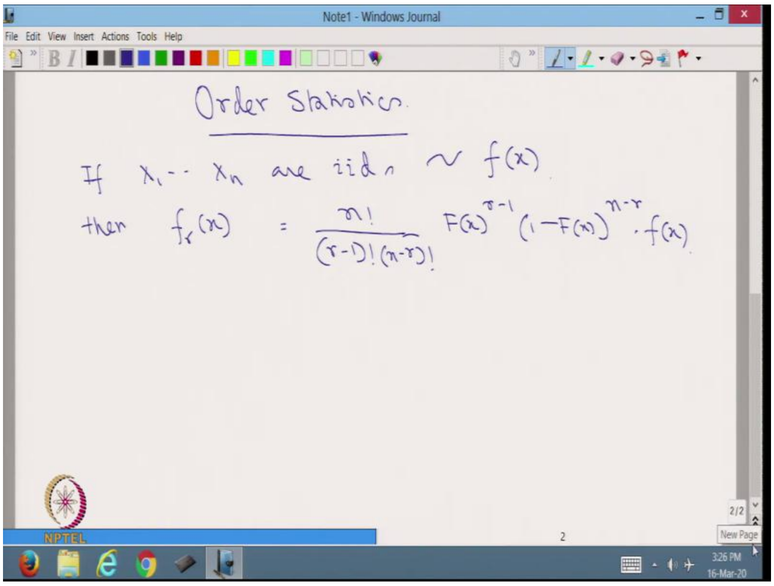
Advanced Probability Theory
Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture 23

(Refer Slide Time: 0:24)



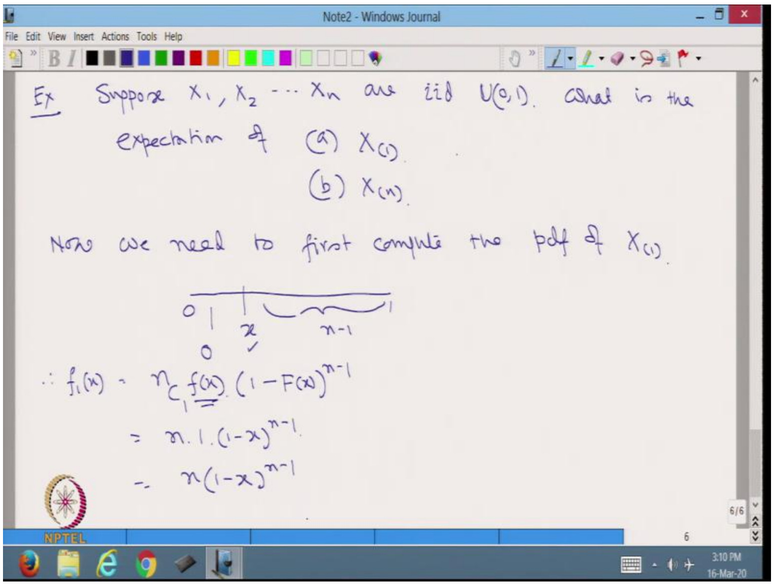
Welcome students to MOOCS series of lectures on Advanced Probability Theory. This is lecture number 23.

(Refer Slide Time: 0:32)



If you remember, we are working on Order Statistics. And we have derived that if X1 X2 Xn are iids following parent distribution f x, then fr x is equal to factorial n upon factorial r minus 1 into n minus r factorial, F x to the power r minus 1, 1 minus F x to the power n minus r multiplied by f x. Before we proceed any further, let me solve some examples.

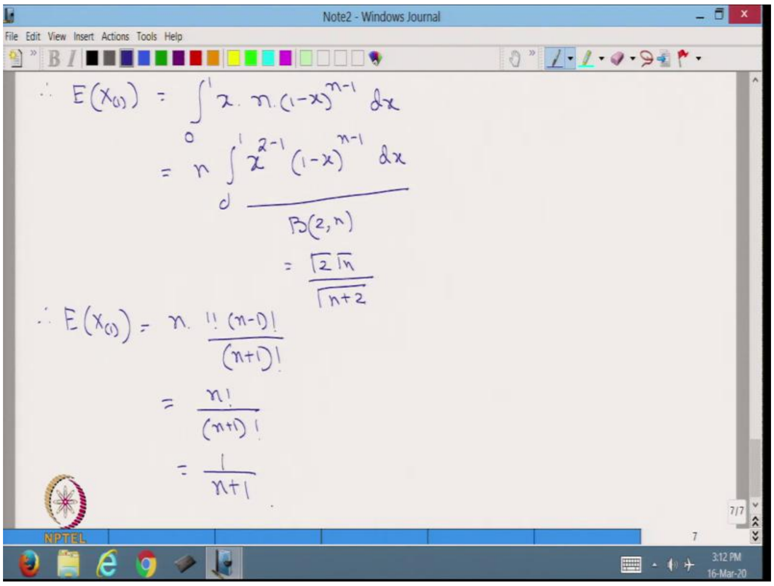
(Refer Slide Time: 1:28)



Suppose X1, X2, Xn are iid uniform 0, 1. What is the expectation of a X1 and b Xn. Now, we need to first compute the pdf of X1. So, we go from the basics 0 1, so we are considering a point x and we are saying that X1 is at x, what is the pdf? So, before x therefore, there are 0 observations, then one observation is that x and remaining n minus 1 observations are in this region.

Therefore, f1 x is equal to nc1, choosing this 1 element, fx multiplied by 1 minus Fx, whole to the power n minus 1 is equal to n, fx is equal to 1 because it is uniform 0 1 and 1 minus Fx is equal to 1 minus x whole to the power n minus 1, is equal to n into 1 minus x whole to the power n minus 1.

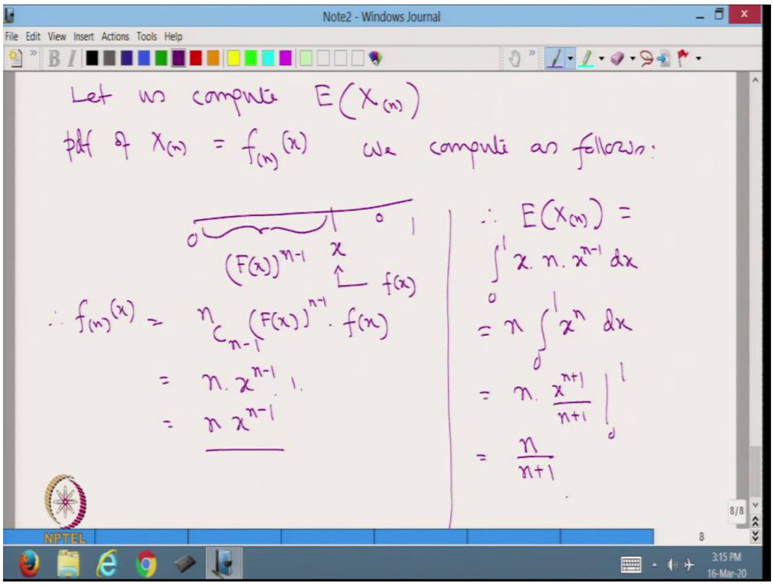
(Refer Slide Time: 3:46)



Therefore, expected value of X1 is equal to integration 0 to 1, x into n into 1 minus x, whole to the power n minus 1, dx is equal to n into integration 0 to 1, x to the power 2 minus 1 into 1 minus x whole to the power n minus 1 dx. Now, this is a beta integral and that is going to give us beta 2, comma n is equal to gamma 2 gamma n upon gamma n plus 2.

Therefore, expected value of X1 is equal to n into gamma 2 is equal to factorial 1, gamma n is equal to factorial n minus 1 gamma n plus 2 is equal to n plus 1 factorial is equal to n factorial upon n plus 1 factorial is equal to 1 upon n plus 1.

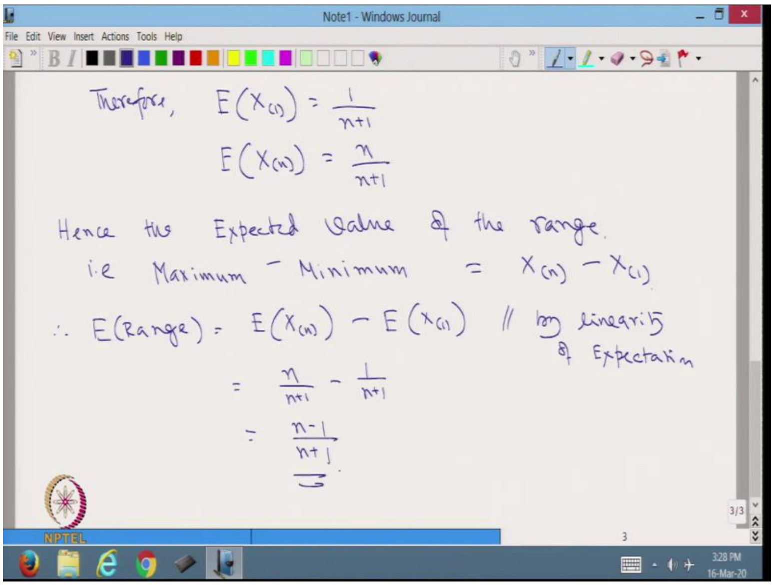
(Refer Slide Time: 5:18)



Let us now compute expected value of Xn the nth other statistic. So, pdf of Xn is equal to fn x, we compute as follows. Suppose, it is 0, it is 1, this is x and nth observation is here, so that will give us fx all other remaining n minus 1 will be in this region. So, that will give us a Fx to the power n minus 1 and there should be 0 observations here.

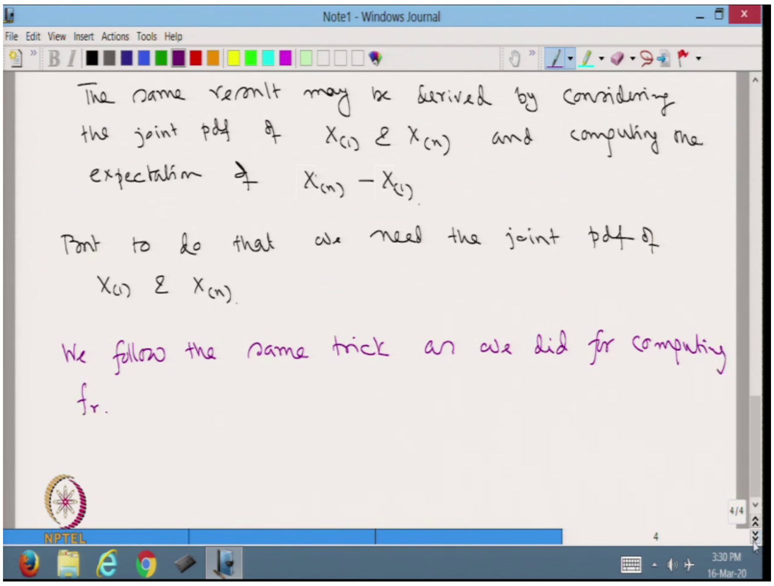
Therefore, fn x is equal to nc n minus 1, Fx to the power n minus 1 into fx is equal to n x to the power n minus 1 into 1 is equal to n x to the power n minus 1. Therefore, expected value of Xn is equal to integration 0 to 1, x n x to the power n minus 1, dx is equal to n into 0 to 1, x to the power n dx is equal to n into x to the power n plus 1, upon n plus 1, 0 to 1, which is is equal to n upon n plus 1.

(Refer Slide Time: 7:26)



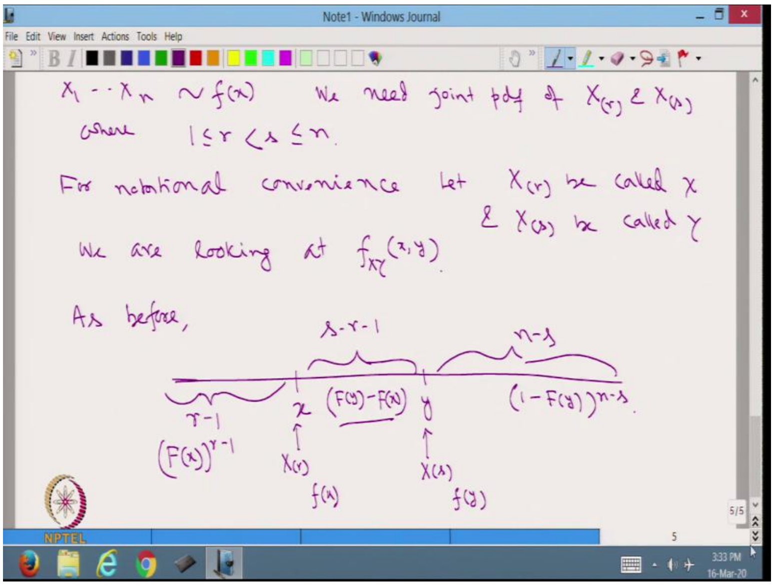
Therefore, expectation of X1 is equal to 1 upon n plus 1, expectation of Xn is equal to n upon n plus 1. Hence, the expected value of the range, what is range? Range is equal to maximum minus minimum is equal to Xn minus X1. Therefore, expected value of range is equal to the expected value of Xn minus expected value of X1 by linearity of expectation is equal to n upon n plus 1 minus 1 upon n plus 1 is equal to n minus 1 upon n plus 1.

(Refer Slide Time: 9:03)



The same result may be derived by considering the joint pdf of X1 and Xn and computing the expectation of Xn minus X1. But to do that we need the joint pdf of X1 and Xn. We follow the same trick as we did for computing fr.

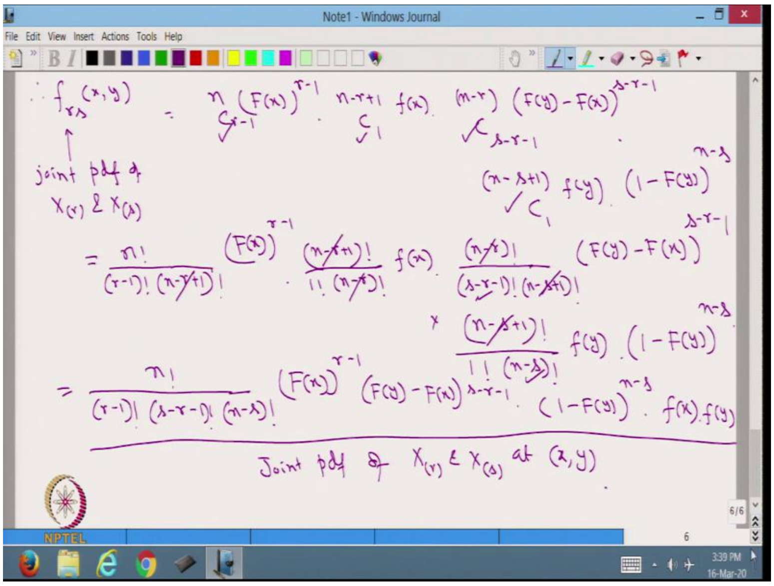
(Refer Slide Time: 10:35)



So, the problem is in X1 X2 Xn follow a pdf fx we need joint pdf of Xr and Xs, where 1 less than equal to r strictly less than s, s less than equal to n. For notational convenience let Xr be called X and Xs be called Y. We are looking at f XY at x,y. So as before, suppose this is the real line, this is the point x, this is the point y, therefore, the rth order statistic is going to take this value which we know will give fx the sth order statistic is going to take the value y and that we get from fy.

Now, r minus 1 many values are here. For them, we shall get Fx to the power r minus 1 s minus r minus 1 values are there. For them we will get Fy minus Fx because they are lying between x and y and for remaining n minus s. We shall get 1 minus Fy to the power n minus s, once we understand that, we can compute the pdf.

(Refer Slide Time: 13:10)

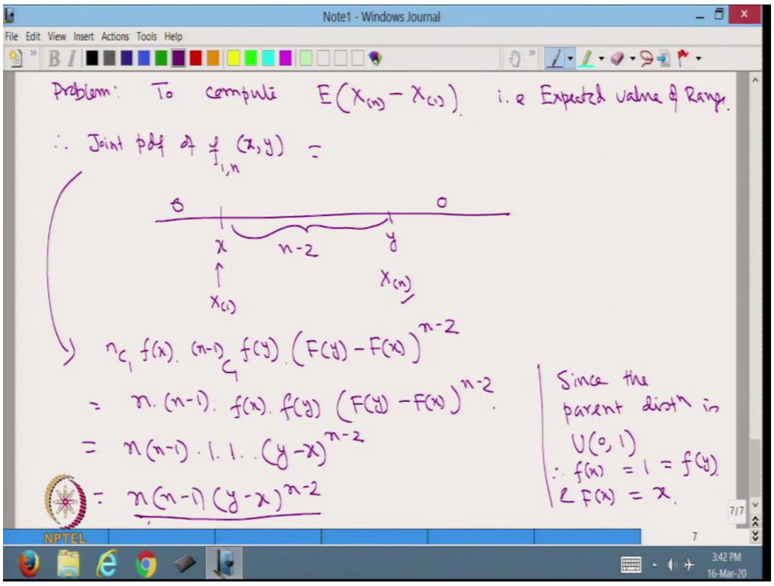


Therefore, f rs at x, y, this is the joint pdf of Xr and Xs, is equal to nc r minus 1, Fx to the power r minus 1, then out of n minus r plus 1, we choose 1 and we give fx out of remaining n minus r, we choose s minus r minus 1, we assign the probability Fy minus Fx to the power s minus r minus 1, then out of n minus s plus 1, we choose 1, give the value fy and the remaining n minus s, they get 1 minus Fy therefore whole to the power n minus s.

Apparently very complicated we try to simplify it, is equal to factorial n upon factorial r minus 1 into factorial n minus r plus 1, Fx to the power r minus 1 multiplied by n minus r plus 1 factorial upon 1 factorial into n minus r factorial and we give the value fx, then n minus r factorial upon s minus r minus 1 factorial into n minus s plus 1 factorial multiplied by Fy minus Fx, Fy minus Fx to the power s minus r minus 1 multiplied by n minus s plus 1 factorial upon 1 factorial into n minus s factorial multiplied by f y multiplied by 1 minus Fy whole to the power n minus s. Is equal to, if we look at, this gets canceled with this, n minus r factorial gets cancelled with n minus r factorial s minus r, n minus s plus 1 factorial gets cancelled with n minus s plus 1 factorial.

Therefore, we get n factorial upon r minus 1 factorial into s minus r minus 1 factorial into n minus s factorial and this multiplied by Fx to the power r minus 1 Fy minus Fx to the power s minus r minus 1 1 minus Fy to the power n minus s into fx into fy. So, that is the joint pdf of Xr and Xs at x, y. We are not going to give the mathematical proof, we have already understood the approach. Therefore, from there we are going to solve the problem.

(Refer Slide Time: 17:38)

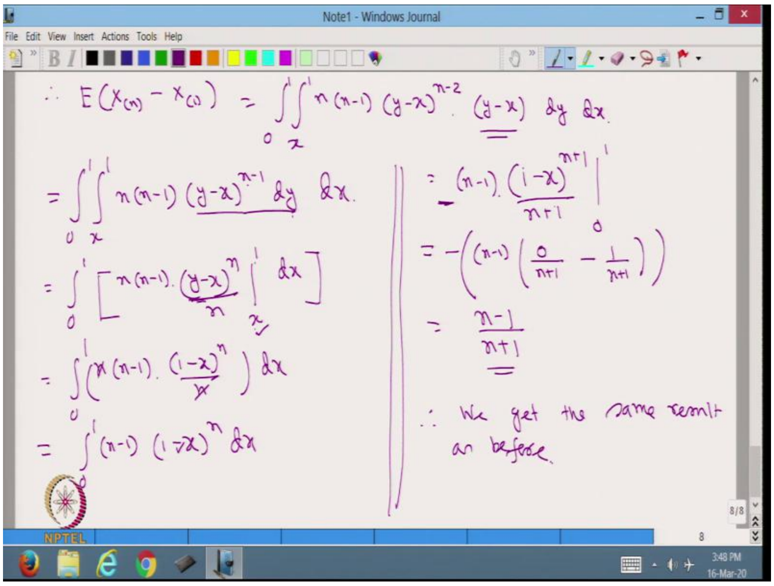


The problem is to compute expected value of Xn minus X1 that is expected value of range. Therefore, what we do, joint pdf of X1 and Xn at the point x, y, therefore we write in a very similar way is equal to this is x X1 is coming here, this is y Xn is coming here. So, there is nothing on this side, there is nothing on this side and remaining n minus 2 observations are here.

Therefore, the joint pdf we can easily write it as nc1 fx out of the n minus 1, n minus 1 c1 that we put at fy multiplied by fy minus fx whole to the power n minus 2 is equal to n into n minus 1 into fx into fy into Fy minus Fx whole to the power n minus 2. Since, the parent distribution is uniform 0, 1.

Therefore, fx is equal to 1 is equal to fy and Fx is equal to x. Hence, putting the value here n into n minus 1 into 1 into 1 into y minus x whole to the power n minus 2 is equal to n into n minus 1 into y minus x, whole to the power n minus 2. So, that is the joint pdf of X1 and Xn.

(Refer Slide Time: 20:20)



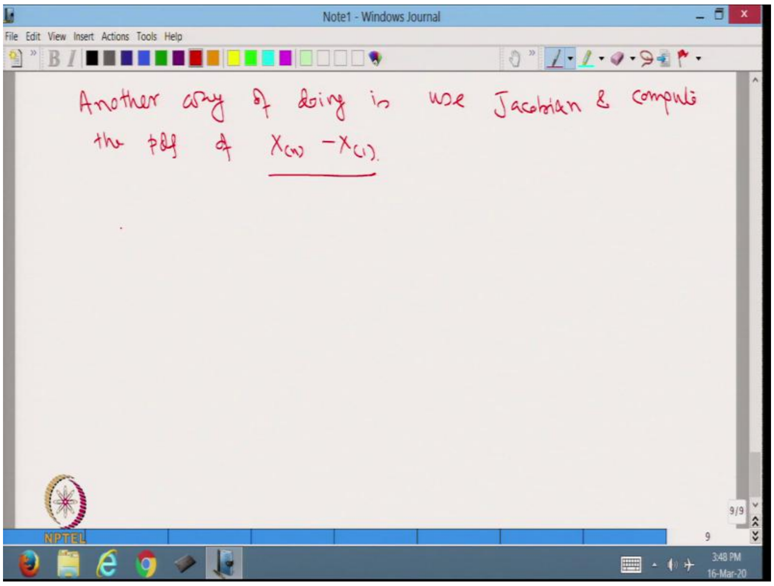
Therefore, expected value of Xn minus X1 is equal to, now we have to use double integration, n into n minus 1 into y minus x whole to the power n minus 2 multiplied by y minus x, so this is the function for which we are trying to find out the expectation into dy dx.

Now, we have to put the range y cannot be less than x, therefore, range of y is equal to x to 1 and x can be anything between 0 to 1 is equal to integration 0 to 1, integration x to 1, n into n minus 1 into y minus x whole to the power n minus 1 dy dx is equal to 0 to 1 n into n minus 1. Now, we are integrating y minus x whole to the power n minus 1 with respect to y.

Therefore, we are getting y minus x whole to the power n upon n, x to 1 dx is equal to integration 0 to 1 n into n minus 1. When we put the value 1 here, we get 1 minus x whole to the power n upon n minus this is going to be 0. If we put x, therefore, we do not need to put that dx is equal to integration 0 to 1 n minus 1 into 1 minus x whole to the power n dx.

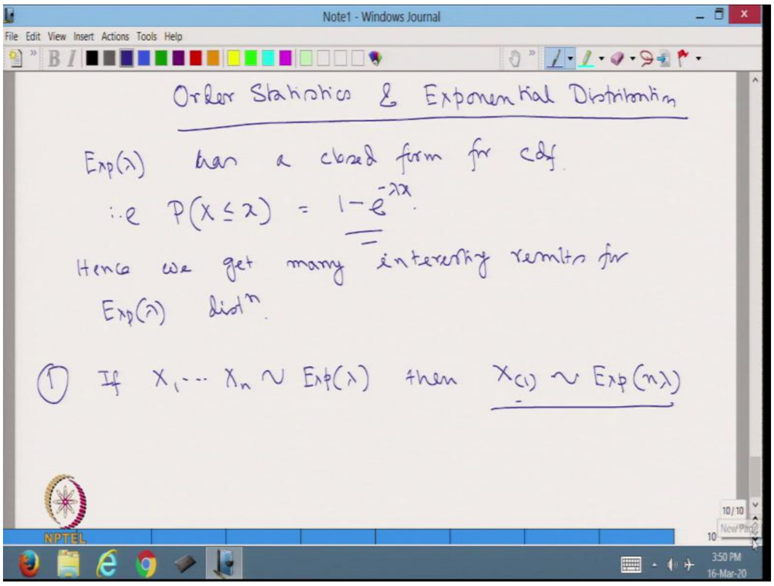
Is equal to n minus 1 goes out, 1 minus x whole to the power n plus 1 upon n plus 1 from 0 to 1, but it comes with a negative sign because x here is negative is equal to minus n minus 1 into 0 upon n plus 1 minus 1 upon n plus 1 is equal to n minus 1 upon n plus 1. Therefore, we get the same result as before.

(Refer Slide Time: 23:50)



Another way of doing it is use Jacobian and compute the pdf of Xn minus X1. I leave that as an exercise.

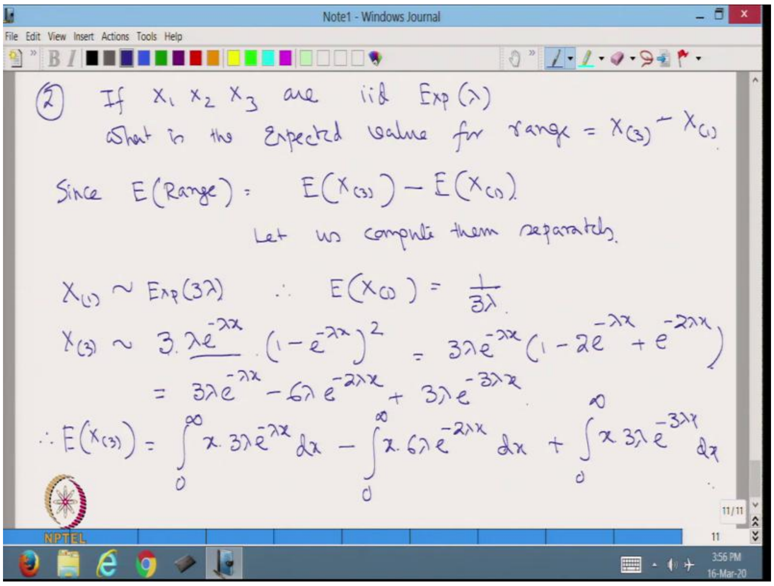
(Refer Slide Time: 24:23)



Let us now look at order statistics and exponential distribution, as I have said before, exponential lambda has a closed form for cdf that is probability X less than or equal to x is equal to 1 minus e to the power minus lambda x. Hence, we get many interesting results for exponential lambda distribution.

1, that we have already seen that, if X1 X2 Xn follow the exponential lambda then X1 the order statistic follows exponential with n lambda, this we have already seen, so I am not going to prove it again. So, let us consider this following result.

(Refer Slide Time: 26:15)

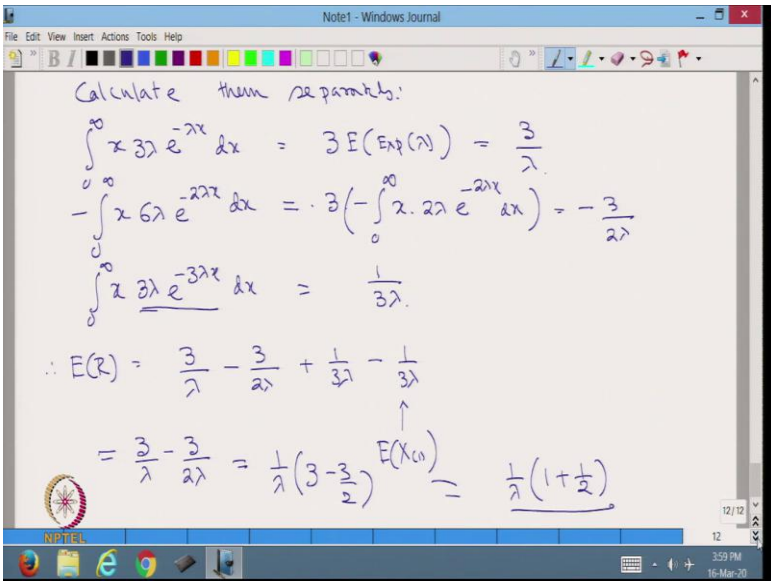


If X1, X2, X3 are iid exponential lambda what is the expected value for range is equal to X3 minus X1. Since, expected value of range is equal to expected value of X3 minus expected value of X1. Let us compute them separately, now X1 follows exponential with 3 lambda that we have seen. Therefore, expected value of X1 is equal to 1 upon 3 lambda. Now, X3 what is the pdf, very simple out of 3, we choose 1 that can be done in 3 ways and put it at the point x that gives us fx and the remaining two they are less than equal to x.

Therefore, 1 minus e to the power minus lambda x whole square is equal to 3 lambda e to the power minus lambda x into minus 2e to the power minus lambda x plus e to the power minus 2 lambda x is equal to 3 lambda e to the power minus lambda x minus 6 lambda e to the power minus 2 lambda x plus 3 lambda e to the power minus 3 lambda x.

Therefore expected value of X3 is equal to integration 0 to infinity x 3 lambda e to the power minus lambda x dx minus 0 to infinity x 6 lambda e to the power minus 2 lambda x dx plus integration 0 to infinity x 3 lambda e to the power minus 3 lambda x dx.

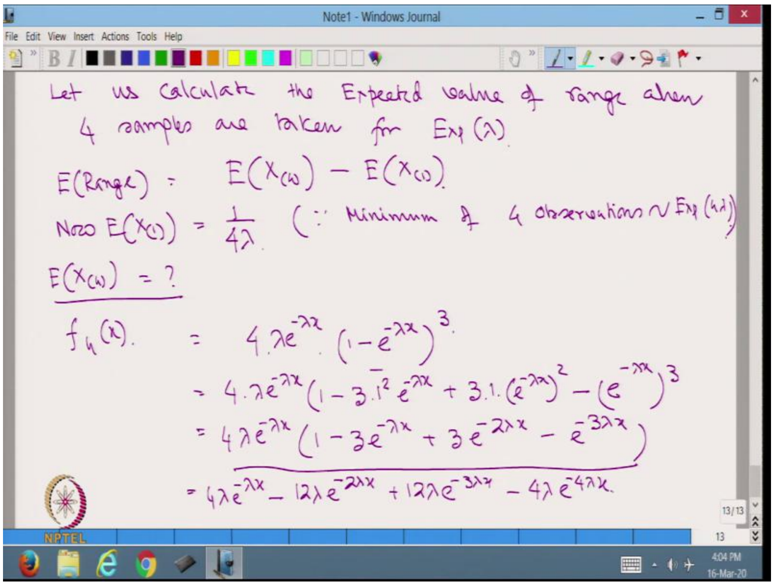
(Refer Slide Time: 29:51)



So, let us calculate them one by one, integration 0 to infinity x 3 lambda e to the power minus lambda x dx is equal to 3 into expected value of exponential lambda is equal to 3 by lambda. The second term is, minus integration of 0 to infinity x 6 lambda e to the power minus 2 lambda x dx is equal to 3 into minus 0 to infinity x to lambda e to the power minus 2 lambda x dx is equal to minus 3 upon 2 lambda and thirdly integration 0 to infinity x 3 lambda e to the power minus 3 lambda x dx is equal to 1 upon 3 lambda, because it is the expectation of exponential distribution with 3 lambda.

Therefore, the expected value of range is equal to 3 by lambda minus 3 by 2 lambda plus 1 by 3 lambda minus 1 by 3 lambda coming from the expected value of X1 is equal to 3 by lambda minus 3 by 2 lambda is equal to 1 by lambda into 3 minus 3 by 2 is equal to 1 by lambda 1 plus 1 by 2.

(Refer Slide Time: 32:24)

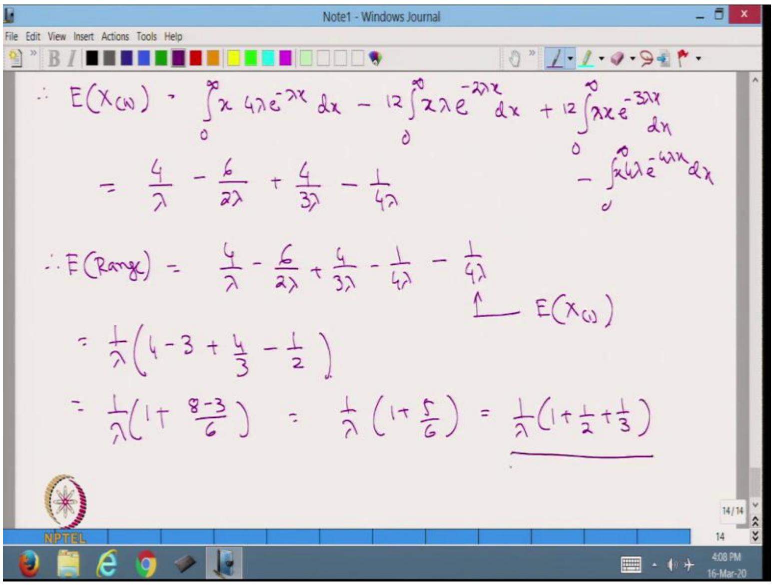


Let us now calculate the expected value of range when 4 samples are taken from exponential with lambda. Now, expectation of range is equal to expected value of fourth order statistic minus expected value of the first order statistic, now expected value of X1 is equal to 1 upon 4 lambda. Since, minimum of 4 observations will follow exponential with 4 lambda.

Now, expectation of X4 is equal to what? To compute that, let us look at f4 of x which is going to be out of 4, you choose 1 and put it at the point x with fx and the remaining 3 will be in less than x. Therefore, fx to the power 3 is equal to 4 lambda e to the power minus lambda x into 1 minus 3 into 1 square into e to the power minus lambda x plus 3 into 1 into e to the power minus lambda x square minus e to the power minus lambda x whole cube.

Is equal to 4 lambda e to the power minus lambda x into 1 minus 3 e to the power minus lambda x plus 3 e to the power minus 2 lambda x minus e to the power minus 3 lambda x, so that is the pdf for the fourth order statistic, we can write it as is equal to 4 lambda e to the power minus lambda x minus 12 lambda e to the power minus 2 lambda x plus 12 lambda e to the power minus 3 lambda x minus 4 lambda e to the power minus 4 lambda x.

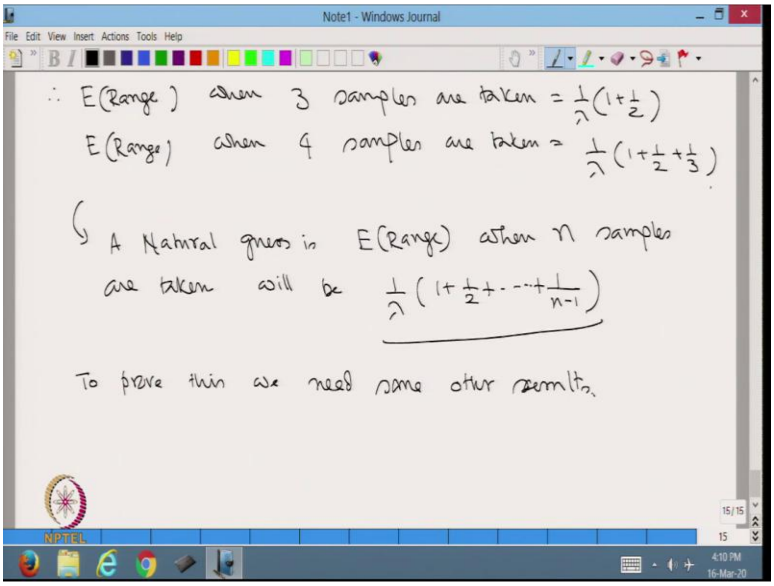
(Refer Slide Time: 36:00)



Therefore, expected value of x 4 is equal to integration 0 to infinity x into 4 lambda e to the power minus lambda x dx minus 12 into 0 to infinity x lambda e to the power minus 2 lambda x dx plus 12 into 0 to infinity lambda x e to the power minus 3 lambda x dx minus integration 0 to infinity 4 lambda e to the power minus 4 lambda x into x dx is equal to 4 by lambda minus, if we take 2 inside, we get 6 upon 2 lambda plus, if we take 3 inside, we get 4 upon 3 lambda minus 1 upon 4 lambda.

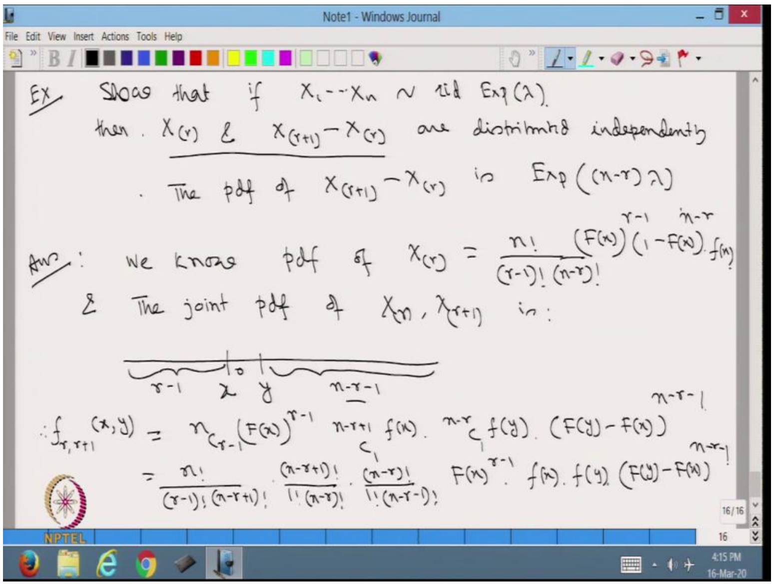
Therefore, expected value of range is equal to 4 by lambda minus 6 by 2 lambda plus 4 by 3 lambda minus 1 by 4 lambda minus 1 by 4 lambda, this is coming from expected value of X1 is equal to 1 by lambda 4 minus 3 plus 4 by 3 minus half is equal to 1 by lambda into 1 plus 8 minus 3 by 6 is equal to 1 by lambda into 1 plus 5 by 6 is equal to 1 by lambda into 1 plus 1 by 2 plus 1 by 3.

(Refer Slide Time: 38:39)



Therefore, expected value of range when 3 samples are taken is equal to 1 by lambda 1 plus half the expected value of range when 4 samples are taken is equal to 1 by lambda into 1 plus half plus 1 by 3. A natural guess is the expected value of range, when n samples are taken, will be 1 by lambda into 1 plus half plus up to 1 upon n minus 1. How to prove it? To prove this, we need some other results, so here is a result.

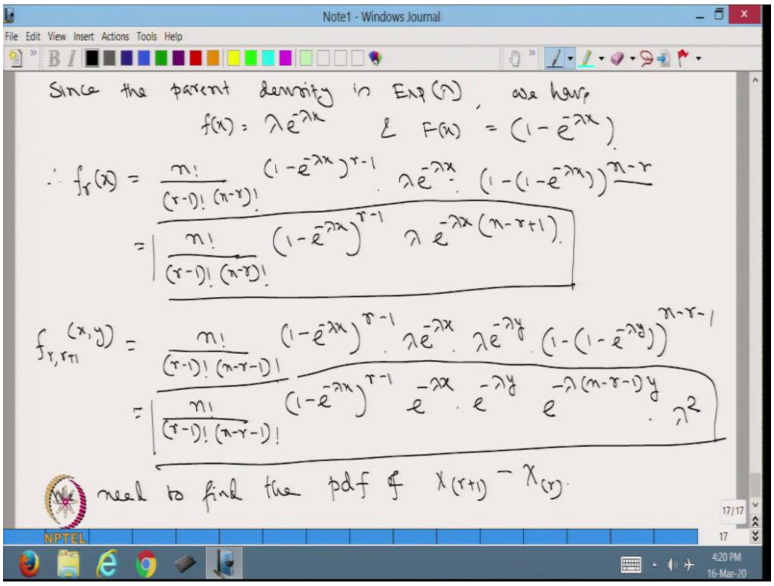
(Refer Slide Time: 40:22)

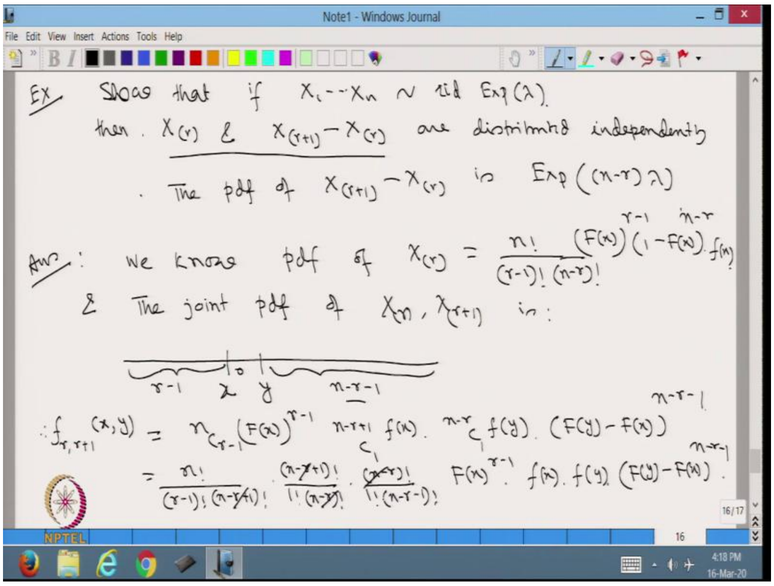


Show that if X1 X2 Xn are iids with exponential lambda, then Xr and Xr plus 1 minus Xr are distributed independently, that is a very interesting result. And second thing is the pdf of Xr plus 1 minus Xr is exponential with n minus r lambda. So, we want to show this answer. We know pdf of Xr is equal to factorial n upon factorial r minus 1 into n minus r factorial Fx to the power r minus 1 1 minus Fx to the power n minus r into fx and the joint pdf of Xr and Xr plus 1 is how to get that, suppose, Xr takes the value x, Xr plus 1 takes the value y, then there are r minus 1 observations here and there are n minus r minus 1 observations are there and 0 observations in between.

Therefore, f r, comma r plus 1 at x, y is equal to ncr minus 1 Fx to the power r minus 1, n minus r plus 1 c1 times fx n minus r c1 times fy and Fy minus Fx whole to the power n minus r minus 1 is equal to after simplification, we will get factorial n upon factorial r minus 1 into factorial n minus r plus 1 into n minus r plus 1 factorial 1 factorial n minus r factorial into n minus r factorial upon 1 factorial n minus r minus 1 factorial multiplied by Fx to the power r minus 1 small fx small fy into Fy minus Fx whole to the power n minus r minus 1.

(Refer Slide Time: 44:39)

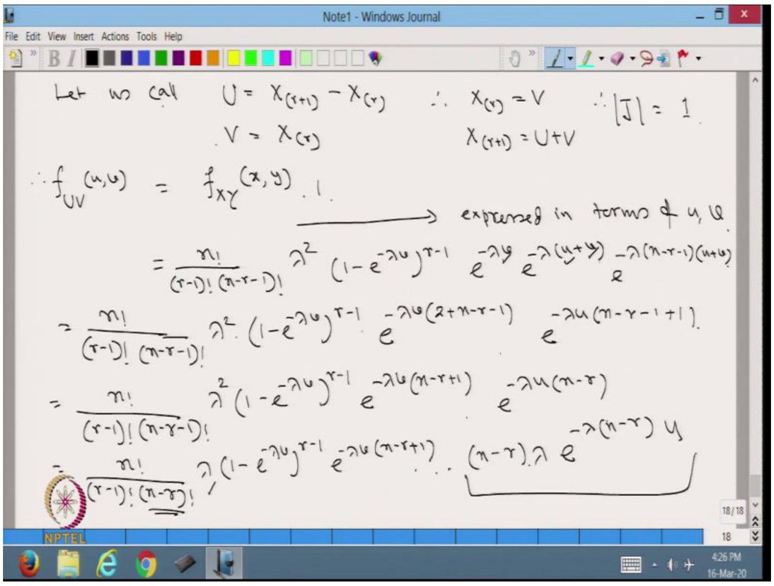




Since, the parent density is exponential with lambda, we have fx is equal to lambda e to the power minus lambda x and capital Fx the cdf is equal to 1 minus e to the power minus lambda x. Therefore, frx is equal to n factorial upon r minus 1 factorial into n minus r factorial 1 minus e to the power minus lambda x to the power r minus 1 lambda e to the power minus lambda x multiplied by 1 minus 1 minus e to the power minus lambda x whole to the power n minus r is equal to n factorial r minus 1 factorial n minus r factorial 1 minus e to the power minus lambda x whole to the power r minus 1 lambda e to the power minus lambda x into n minus r plus 1. So, that is the pdf of the rth order statistic.

What is a fr, comma r plus 1 the joint density at x, y by using the formula we derived just now, in which we can cancel this we have n factorial upon r minus 1 factorial into n minus r minus 1 factorial into 1 minus e to the power minus lambda x whole to the power r minus 1 lambda e to the power minus lambda x lambda e to the power minus lambda y into 1 minus 1 minus e to the power minus lambda y whole to the power n minus r minus 1 is equal to n factorial r minus 1 factorial n minus r minus 1 factorial 1 minus e to the power minus lambda x to the power r minus 1, e to the power minus lambda x into e to the power minus lambda y into e to the power minus lambda n minus r minus 1 y multiplied by lambda square. So, that is the joint pdf of x, y. We need to find the pdf of Xr plus 1 minus Xr.

(Refer Slide Time: 48:20)



So, let us call U is equal to Xr plus 1 minus Xr and V is equal to Xr. Therefore, Xr is equal to V, Xr plus 1 is equal to U plus V. Therefore, Jacobean as we have computed many times is equal to 1. Therefore, f UV at u, v is equal to joint pdf of Xr and Xr plus 1 at the point x, y multiplied by 1 expressed in terms of u and v.

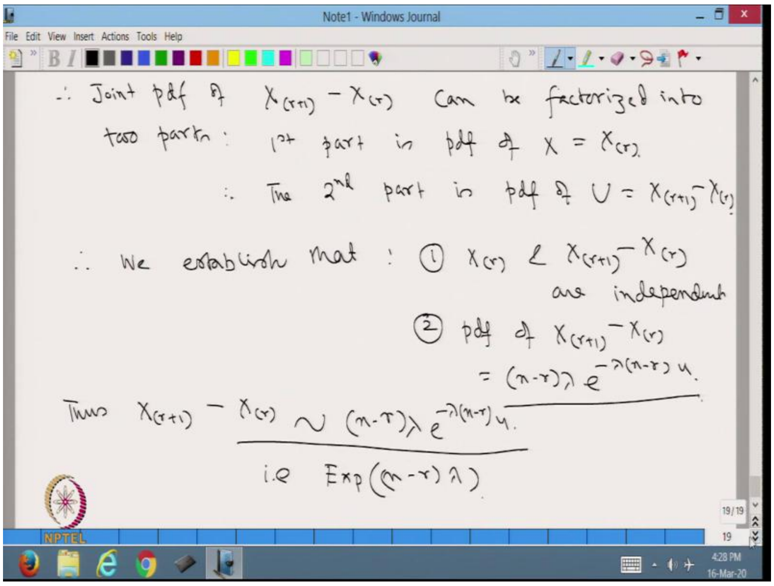
Is equal to n factorial upon r minus 1 factorial n minus r minus 1 factorial, lambda square 1 minus e to the power minus lambda v whole to the power r minus 1, e to the power minus lambda v, e to the power minus lambda u plus v, e to the power minus lambda n minus r minus 1 into u plus v multiplied by 1 so that is not needed.

Is equal to n factorial r minus 1 factorial into n minus r minus 1 factorial lambda square 1 minus e to the power minus lambda v whole to the power r minus 1 e to the power minus lambda v into 1 plus 2 plus n minus r minus 1 multiplied by e to the power minus lambda u into n minus r minus 1 plus 1.

Is equal to n factorial r minus 1 factorial n minus r minus 1 factorial lambda 1 minus e to the power minus lambda v whole to the power r minus 1 e to the power minus lambda v to the power n minus r plus 1 multiplied by e to the power minus lambda u whole to the power n minus r.

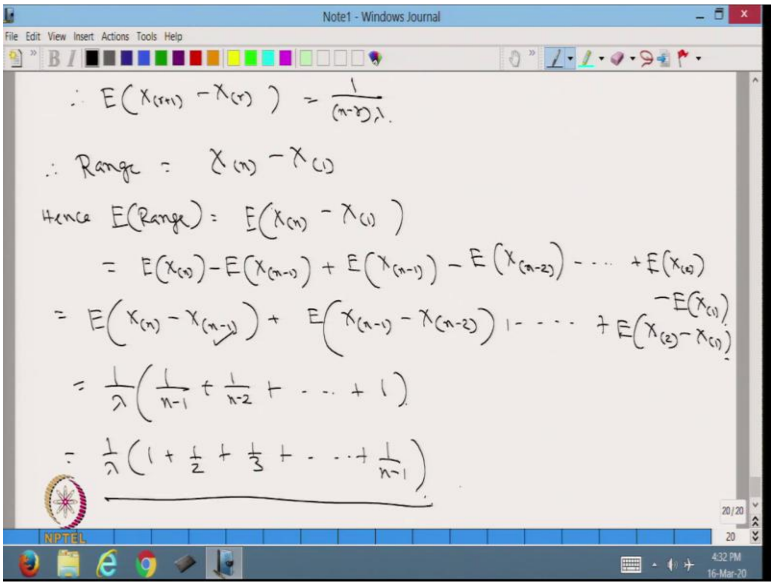
Is equal to n factorial r minus 1 factorial n minus r factorial lambda 1 minus e to the power minus lambda v to the power r minus 1 e to the power minus lambda v to the power n minus r plus 1 into, since we have multiplied by n minus r in the denominator, we take out that n minus r, then one of the two lambdas, one we have come here and the other will come here multiplied by e to the power minus lambda n minus r times u.

(Refer Slide Time: 52:44)



Therefore, joint pdf of Xr plus 1 minus Xr can be factorized into two parts. The first part is pdf of X, which is the rth order statistic. Therefore, the second part is pdf of U is equal to Xr plus 1 minus Xr. Therefore, we establish that 1, Xr and Xr plus 1 minus Xr are independent and secondly, pdf of Xr plus 1 minus Xr is equal to n minus r, lambda e to the power minus lambda n minus r u. Thus, Xr plus 1 minus Xr is distributed as n minus r lambda e to the power minus lambda n minus r u, that is exponential with n minus r lambda.

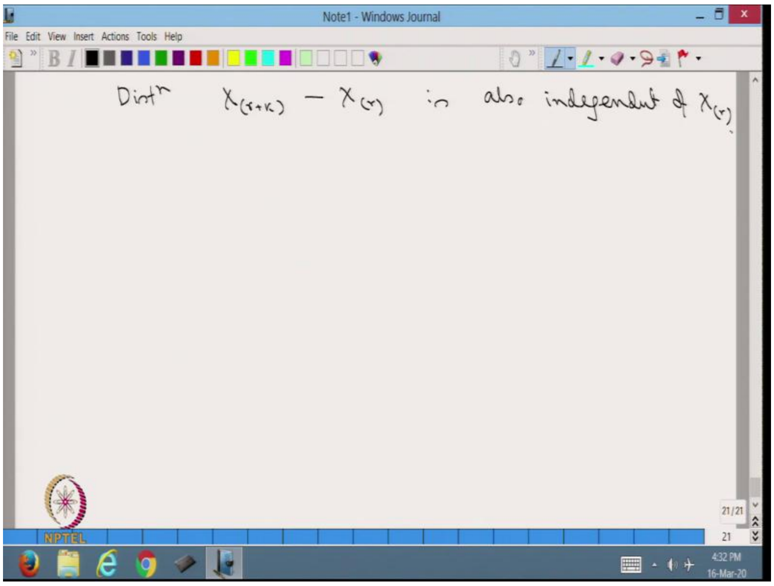
(Refer Slide Time: 55:00)



Therefore, Xr plus 1 minus Xr is equal to 1 upon n minus r lambda. Therefore, range is equal to Xn minus X1. Hence, expected value of range is equal to expected value of Xn minus X1 is equal to expected value of Xn minus expected value of Xn minus 1 plus expected value of Xn minus 1 minus expected value of Xn minus 2 and like that, if we go, we get plus expected value of X2 minus expected value of X1.

Is equal to expected value of Xn minus Xn minus 1 plus expected value of Xn minus 1 minus Xn minus 2 like that, expectation of X2 minus X1 is equal to 1 upon lambda 1 upon n minus 1 by taking r is equal to n minus 1 plus 1 upon n minus 2 plus up to 1 is equal to 1 upon re lambda 1 plus half plus 1 by third 1 by 3 up to 1 upon n minus 1.

(Refer Slide Time: 57:30)



So, that is the result that we wanted to prove okay friends, these are very important result. In fact, the result can be even more generalized, that it is not only successive difference even expectation of, even the distribution of Xr plus k minus Xr is also independent of Xr. I like you to prove that as an exercise. Okay friends, I stop here today, from the next class I shall start with convergence theorems for probability distributions. Okay friends, thank you.