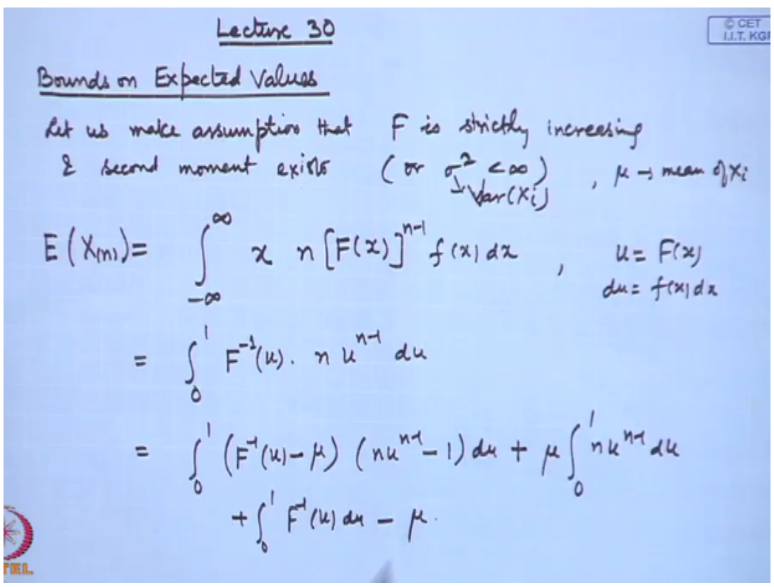
Statistical Methods for Scientists and Engineers

Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology – Kharagpur

Lecture – 30
Non parametric Methods - III

As, I mentioned in the last lecture that were calculating the expected values of the; I gave expressions for the expectations or the variance etc., of the order statistics, when the sampling is done from uniform distribution. I also mentioned that it is quite complicated to obtain the same expressions for the general distribution, if the sample is from a general continuous population. The main reason is that the distribution of the order statistics involves the powers of F and also the powers of 1 – F.

(Refer Slide Time: 01:29)



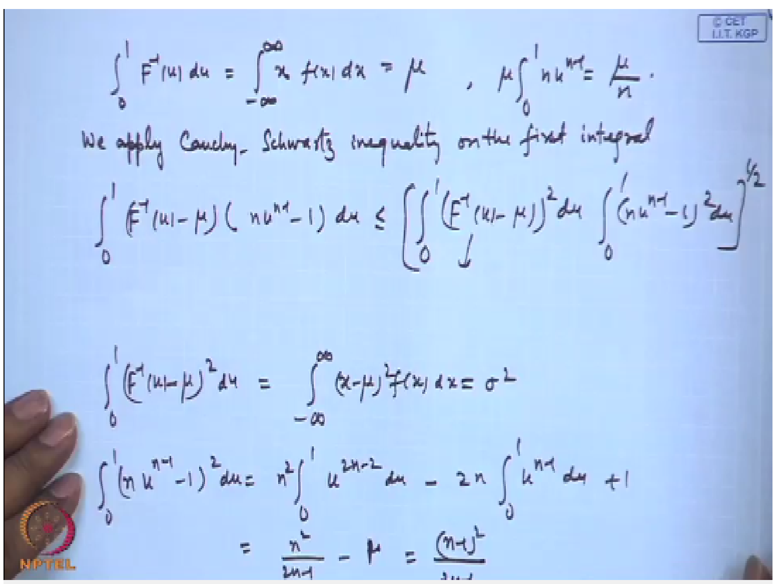
And therefore, for a typical distribution say, normal distribution etc., these expressions will be intractable that means we cannot evaluate them in the closed form. Therefore, certain procedures have been developed for which we can at least obtain approximations or we can obtain bounds, so first I will discuss the method of lower or the upper bounds for the expected values. So, first assumption is that let us make assumption that F is strictly increasing.

And second is that second moment exists, so in general or we can say, sigma square is finite that means the variance of say, Xi that exists. Under these assumptions, these expressions will be obtained. So, let us consider say, for example expectation of the largest, now we have derived the distribution of the largest as n times F of x to the power n -1 fx and multiplied by x here, now we substitute say, u = Fx.

So, of course du = fx dx, then this is becoming = integral from 0 to 1 and place of x, we will have F inverse, this is the reason why we made the assumption that effort is to be increasing because then I can ride on the inverse function in a unique question that F is strictly increasing because then, I can write down the inverse function in a unique fashion. Then, this is becoming u to the power n – 1 du.

Now, this terms we make some adjustment here, we write it as 0 to 1, F inverse u – Mu. Now, what is Mu? Mu we take to be mean of Xi, so then you consider F inverse u – Mu, n u to the power n -1, -1 du that means I have added and subtracted some terms here; n times u to the power n -1 du + 0 to 1 F inverse u du – Mu. Now, this is a cross product on this I apply Cauchy Schwartz in equality and also let us look at what are these terms.

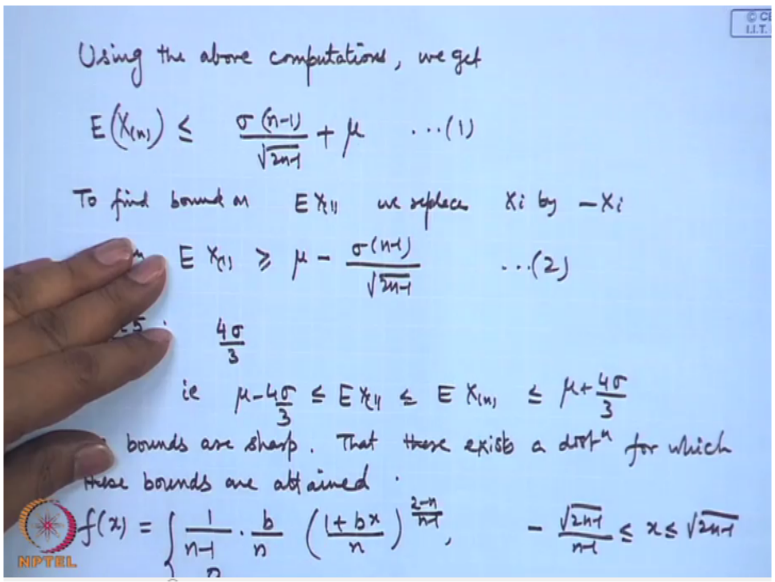
(Refer Slide Time: 04:04)



Let us check each of them separately, so if I look at 0 to 1 F inverse u du, this is nothing but integral x fx dx because if I put F inverse u is = x, then I will get Fx is = u, then du is becoming fx dx. So, this is nothing but Mu itself. Similarly, the integral 0 to 1 nu to the power n -1 that is simply becoming Mu/n, then we apply Cauchy Schwartz equality on the first integral that means on this term, I apply the Cauchy Schwartz equality.

So, this will give me that is 0 to 1 F inverse u – Mu nu to the power n -1 -1 du that is <= integral 0 to 1 F inverse u - Mu whole square du integral 0 to 1 nu to the power n -1-1 whole square du, whole to the power 1/2/ Once again, what are these terms? If I look at this term here 0 to 1 F inverse u – Mu, if I put F inverse u = x, then this is becoming - infinity to infinity; x - Mu square fx dx that is = sigma square.

(Refer Slide Time: 06:40)



And the second term can be easily evaluated that is 0 to 1 n times u to the power n -1-1 square du, so that is = 0 to 1 n square u to the power 2n -2 du -2n, u to the power n – 1du + that is = n square/ 2n -1-2+1, so that is -1 that is = n -1 whole square/ 2n -1. So, let me substitute all the values here, so we get, using the above computations, we get; we started with the term expectation of Xn, so the expectation of Xn is <= sigma * n-1/ root 2n -1 and then what are the other terms that we are getting here?

This term is giving me Mu, this is giving me Mu, this is Mu, so this is getting cancelled out, so I get only Mu here, so you can say this is a bound here. To find bound on expectation of X1, we replace Xi/ - Xi, so then we get expectation of X1 >= Mu - sigma n -1/ root 2n – 1, so basically you can see that moments of the order statistics lie between Mu - sigma n – 1/ root 2n – 1 to Mu

+ sigma * n – 1/ root 2n – 1. Let us take say, some special case for example, I take n = say 5, okay.

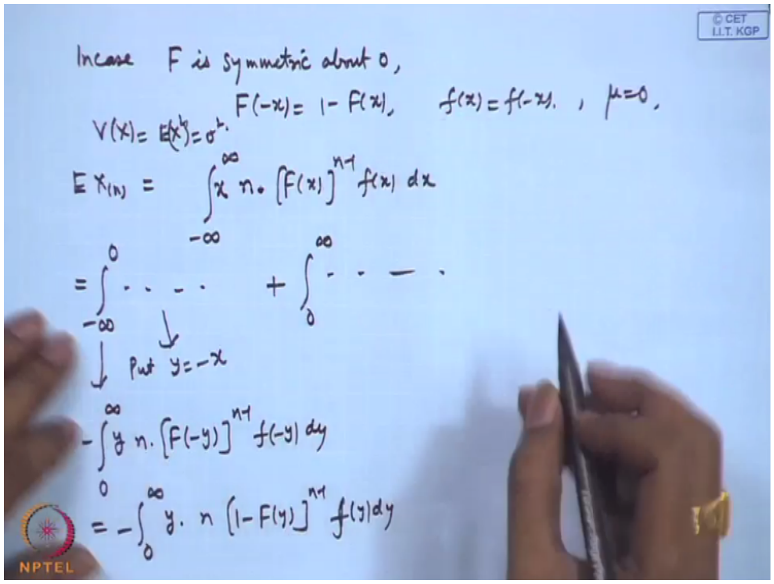
If I take n = 5, then what value I will get? This sigma n -1 that is becoming 4 and this will become 3, so 4/3 that means we are getting expectation of X1 is >= Mu -4 sigma/3 <= Mu +4 sigma/3, so it is helpful in that sense for example, we know expectation of each Xi is M, so now I want to talk about expectation of Xn, so the upper bound that is attained when I have a sample of size 5 is Mu +4/ 3 times the standard deviation of the random variable.

And similarly, the lower bound will be Mu -4/3 times of the standard, so in some sense this is like giving you a scale for example, when we were dealing with the normal distribution, then we consider like sigma limits Mu + - sigma limits Mu + - 2 sigma limits are Mu + -3 Sigma limits etc, so this is giving us the similar bound in the case of the order statistics from any distribution. Then, 1 question may arise for example, this television that I have done is through Cauchy Schwartz equality.

Now, we know that in the Cauchy Schwartz equality, the equality is attained, if the 2 functions F and G are basically linearly related. Now, if we apply that condition and certainly equality must be attained, so we can actually show that this bounds are sharp that is; there exists a distribution for which these bounds are attained. Actually, that example is given here, so let me mention that thing here.

That is if I consider fx = 1/ n – 1, b/n, 1 + bx/ n to the power 2-n/n -1, where x is from root 2n -1/ n-1 to root 2n -1. If I consider this density of course, it is 0 outside this range. For this density, both the bounds are attained, which can be easily checked I am not giving the calculations here, actually if we consider this the form of the CDF is also in a closed form and therefore all the calculations can be easily done here.

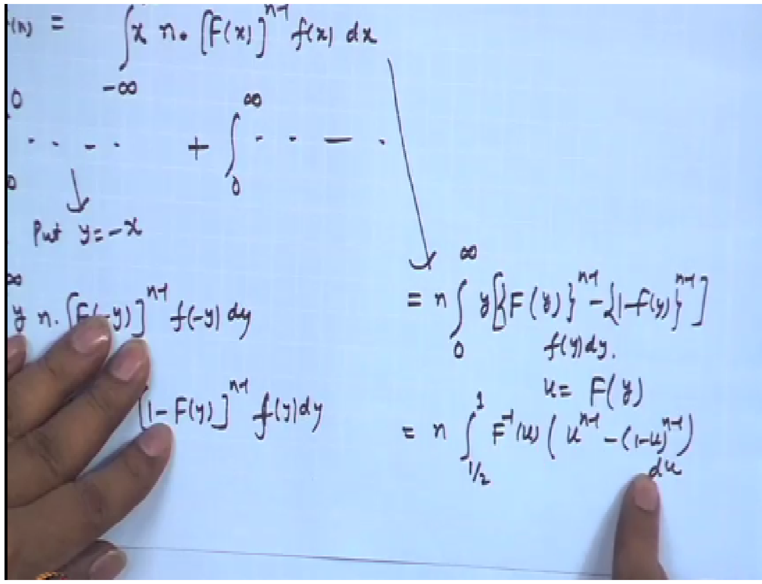
(Refer Slide Time: 11:46)



Now, if the distribution is symmetric about 0, then that means Mu is 0 in that case, there is a further improvement over the bound, let me give it here, in case F is symmetric about 0, then let us consider. Generally, if it is symmetric, then we have this property that F of – x = 1 – F of x and also small f of x is = f of –x and you will have Mu is = 0 that means variance will be = expectation of x square.

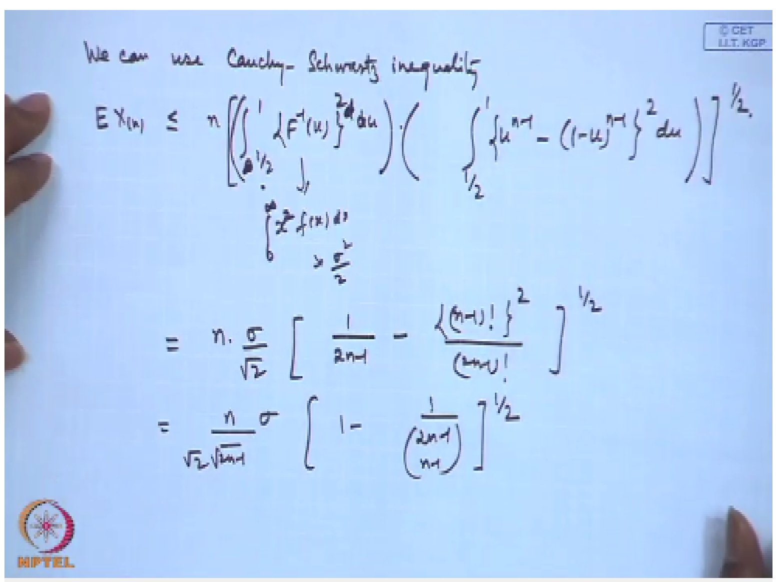
Now, in this case let us consider say expectation of Xn, so that is = - infinity to infinity n times F of x to the power n -1 fx dx. Now, this is multiplied by Xe, I split it into that 2 region – infinity to 0 +0 to infinity. In the first one, you put say, y = -x, then this is becoming 0 to infinity, then, X is becoming –y, then you get n times F of –y to the power n – 1, f of –y dy. Now, using this property of symmetry here, this is becoming = -0 to infinity yn 1 - Fy to the power n – 1, f of y, dy.

(Refer Slide Time: 13:40)



So, again you can put y = x, so I can now consider this integral is 0 to infinity, this integral is 0 to infinity, this is with a - sign and here I have ny that is nx here and fx that is common here, so this term then becomes = this term can be then written as 0 to infinity n times y F of y to the power n -1-1 - fy to the power n -1 fy dy. Now, you can again put as before u is = Fy, if we do that, then this is simply becoming = n times 1/2 to 1.

(Refer Slide Time: 14:58)



Why 1/2 to 1? Because F inverse 0, since it is symmetric about 0, then F of 0 = 1/2, so F inverse 0 =; F inverse 0 will become = 1/2, so this is now then F inverse u and then you get u to the power n -1-1 –u to the power n -1 du. Again, you see this is interesting here this is coming as a product of 2 terms, so I can apply Cauchy Schwartz equality, we can use Cauchy Schwartz, so you will get expectation of Xn <= n times integral.

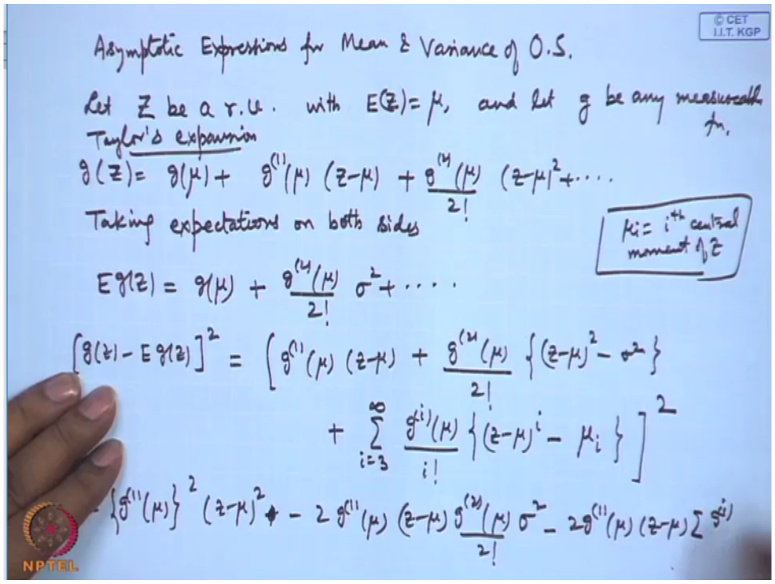
Well, it is 1/2 to 1 F inverse u square du * integral 1/2 to 1 u to the power n -1-1 -u to the power n -1 is square du, whole to the power 1/2. So, it is a matter of calculations here, this is actually becoming x square fx dx from 0 to infinity. So, since I have assumed that the mean is 0 and it is a symmetric function, so it is actually = - infinity to infinity 1/2 of that, so this is basically = sigma square/ 2.

This can also be easily evaluated, so I just write down the values here, it is becoming = n sigma/ root 2 because power 1/2 is coming here and in the second part it is coming 1/ 2n – 1 – n-1 factorial is square/ 2n -1 whole to the power 1/2. Of course, you can further simplify it as = n/ root 2n – 1 * root 2 sigma 1- 1/2n – 1; see n - 1 to the power 1/2. If you compare with the bound that we obtained here, it was sigma n -1/ root 2n -1 + Mu.

So, this Mu is actually becoming = 0 here; here you can see that sigma this; divided by root 2n -1 is coming here but this is n -1 here, that is becoming n here and here you are getting root 2, so this bound is much smaller that means it is a sharper bound, if we are having additional information that the distribution is symmetric about 0, okay. So, here what you can see here that we have made use of Cauchy Schwartz equality for the evaluation.

Then of course, one may question that why did I am not consider rth here, if you consider the rth one, then there is a difficulty here. See in the rth one, 2 terms will be coming here, so here it is easy because if I do it for one of them, then I am able to convert it into x but if I do the second one, then it is becoming 1-x and I would not be able to evaluate there, I mean it will become a much more complicated.

(Refer Slide Time: 18:43)



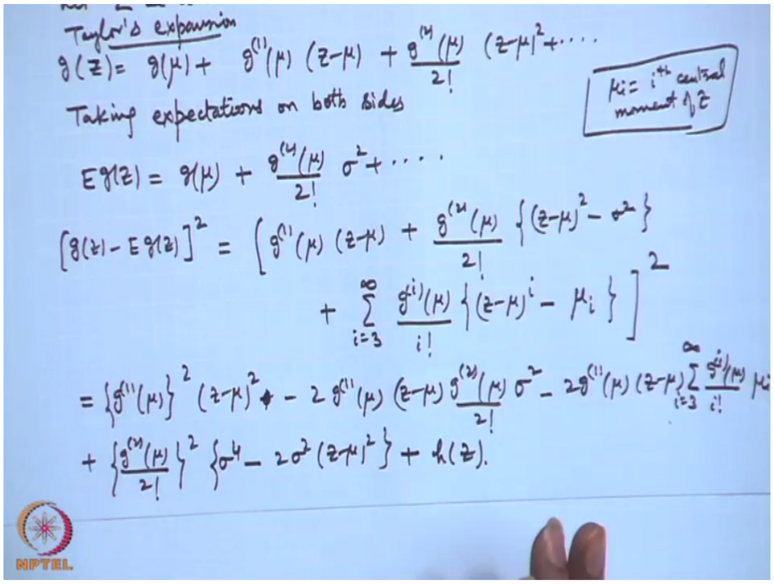
Then the second is the asymptotic here, so let me talk about that. So, we look at asymptotic expressions for mean and variance of order statistics. Firstly, I develop a general formula, so let us consider say, let z be a random variable with expectation of z giving to be say, Mu and let g be any measurable function, so gz, let us consider here that is = g Mu + g1 Mu, z – Mu + z2 Mu/ 2 factorial, z - Mu square and so on.

I am applying Taylor’s theorem here; Taylor's expansion, so I am assuming that g is a nice function that means it is differentiable many times, at least 2 times we are writing down the term here, so if I take expectations on both the sides; taking expectations on both sides, I get expectation of gz that is = g Mu, now in the second one, if I take expectation, this is simply vanishing and then I get here expectation of z - Mu square.

So, I have assumed here expectation of this; this is becoming variance here, so g2 Mu/ 2 factorial sigma square and so on. So, if I consider say, gz - expectation of gz, then this first term will get cancelled out, so let me take a square here, then second term will be coming here that is g1 Mu z - Mu + in the second one, I will get g2 Mu/ 2 factorial and here I get z - Mu square – sigma square.

There are other terms also, we can write it as say, sigma gi Mu/ i factorial z - Mu to the power i -; basically Mu i that is the; Mu i is the ith central moment of z, if I take this, then I get this term here and then this is from i = 3 to infinity is square of this. So, if I expand this, I will get g1 Mu square z - Mu square, then next term will give me -2 times; see this sigma square term is there, so if I consider this expansion that is sigma square, the cross product with this.

(Refer Slide Time: 23:15)

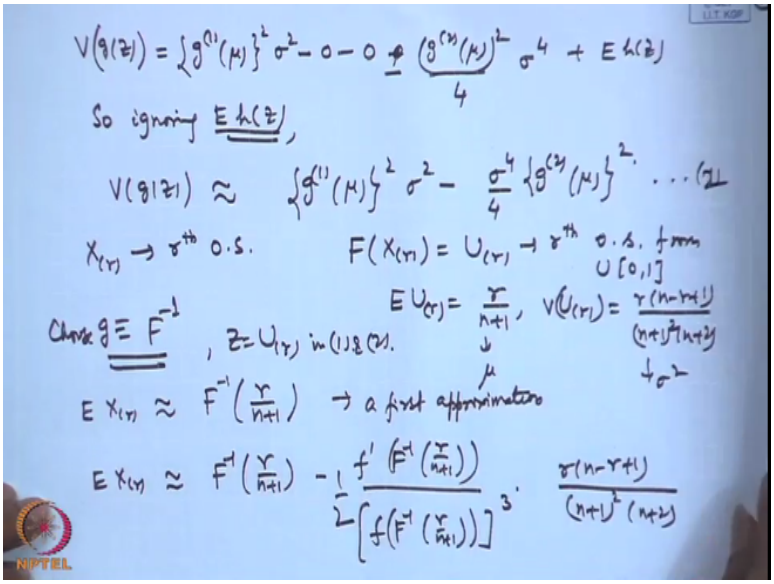


I will get 2 times g1 Mu z – Mu, then g2 Mu/ 2 factorial and sigma square and then certainly, all the terms that I will be obtaining from this that is when I square root it, then I will get the cross product with this one, so let me just put that general form here, -2 g1 Mu z - Mu sigma gi Mu/ i factorial * Mu i; i = 3 to infinity, all these terms will be coming there, then I will get square of this term that is + g2 Mu/ 2 factorial is square.

And then I will get z -; because I have already taken the cross product of this separately, so I will get a term, which is giving me sigma square; so sigma to the power 4 basically and then cross product of this with this itself, so I will get sigma to the power 4 - twice sigma square z - Mu square, so basically I am just arranging in a; you can see that I am arranging the terms in a particular way.

Then, of course then there will be terms with u square this, then that will give me z – Mu to the power 4. Here, the powers are starting from cube, so when I square it, then in the cross product with this itself, I will get 3, 4 and so on. So, all those terms I write as h of z, so okay what I have done, I have written the terms only up to square order of z – Mu, now that was the reason that I assume up to the variance only because I am not going beyond that.

(Refer Slide Time: 25:07)



This approximation is dependent upon this choice here. Now, let us consider here, variance of gz that means if I take expectation here that will give me the variance of gz, it is equal to expectation of these terms. If I take expectation here, this will give me sigma square, here this term will be simply vanished, this term will simply vanish, this term will again give me sigma square.

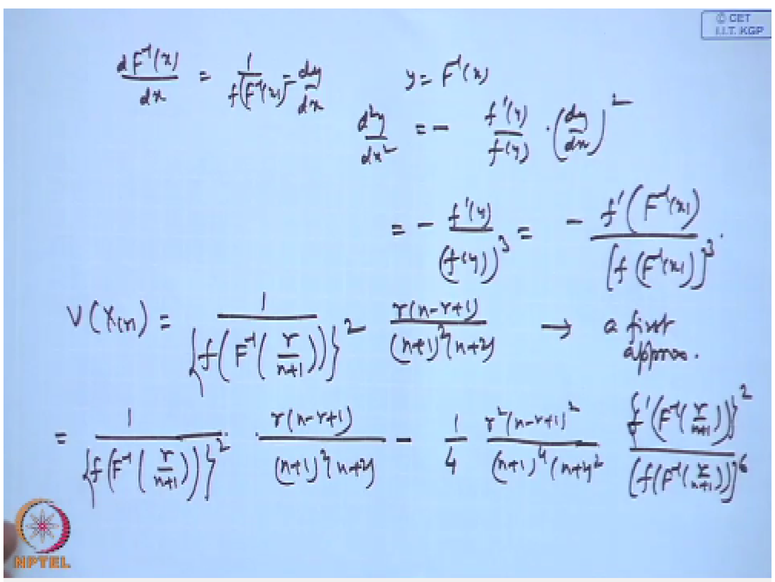
So, sigma to the power 4 -2 sigma to the power 4 + expectation of hz, so this we write as g1 Mu is square sigma square – 0 - 0+ g2 Mu square/4 - sigma to the power 4 + expectation of hz. So, if we ignore these terms; ignoring expectation of hz, variance of gz is approximately = g1 Mu is square sigma square - sigma to the power 4/4 g2 Mu square, this gives an approximate expression for the variance of any function of a random variable.

In terms of Mu and sigma square, so that means the first and second moment must be of a known form. Now, what we do; we make use of the order statistics here, that is; since if I am considering order statistics, then f of that is becoming uniform. So, if you use that, then let us see, so this Xr is the r th order statistics okay, so if I consider F function; F function is giving me Ur that is the r th order statistics from uniform 0, 1.

Why am taking this because I know the mean etc of this, so actually I am putting x is = Ur that is basically we are having expectation of Ur = r/ n +1 and variance of Ur is also known that is r * n – r + 1/ n + 1 square * n + 2, so we will actually; so this Mu 1 sigma square that is this is Mu and this is sigma square and F inverse if I put here that is g function I am taking to be; choose g = F inverse, okay.

So, if you do that and z = Ur in this formula, let me call it 1, then what I am getting? I am getting here; well, let us consider the first expression in 1 and this expression that I wrote here, I can call this as 1 and this as 2 say, so in 1 and 2, so what I will get here that expectation of Xr is assumed totally equivalent to F inverse r/n + 1, so this you can consider as a first approximation and I can also consider 2 terms there.

(Refer Slide Time: 29:24)



F inverse r/ n + 1 – f prime F inverse r/n + 1/ F of F inverse r/n + 1 whole cube; r * n – r + 1/ n + 1 square * n +2, there will be 1/2 coming here, actually this term is coming from the derivative of F inverse there because let me show the calculation here for the derivative. See, if I am conceding say d of F inverse x/ dx, then that is actually becoming 1/f F inverse x and if I considers say, for example if I write y = F inverse x.

Then, what is the value of say, s2y/dx square; see this is actually dy/dx, okay. So, if I consider second derivative, I will get - f prime y/ fy dy/dx square, so this I can write it as –f prime y/fy by whole cube that is = - f prime of F inverse x/ f of F inverse x whole cube. Similarly, if I want to consider say, variance of Xr, then I will get second order term here by using this one as 1/ f F inverse r/n + 1 whole square that is this term here.

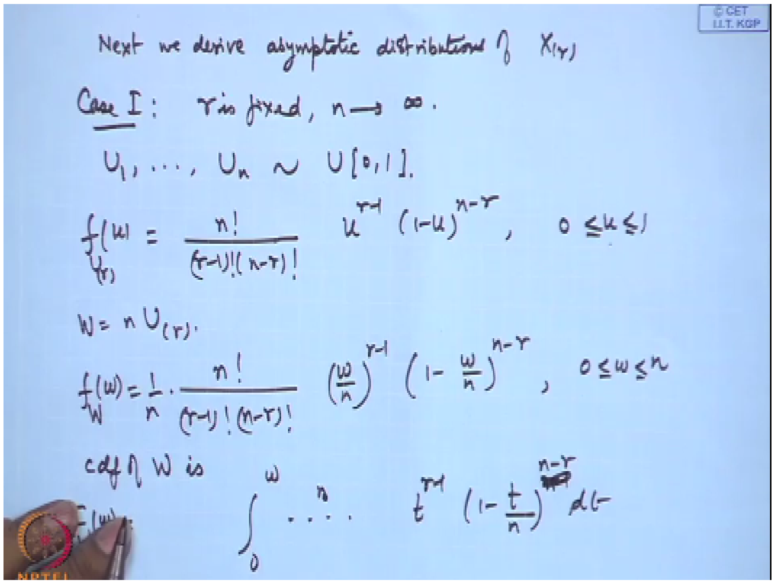
The first order derivative is square here multiplied by the variance that is r * n – r + 1/ n + 1 square * n + 2, so this is actually a first approximation. I can consider 2 terms also that is if I take this one, then it will become second approximation will be 1/ f F inverse r/ n +1 is square, r * n – r + 1/ n + 1 square * n = 2 and second term will give me 1/4 that is this one; sigma to the power 4, so I will get r square * n – r + 1sqaure/ n + 4 to the power 4 * n + 2 square.

And then this f prime F inverse r/n + 1 is square divided by f of F inverse r/ n +1 whole to the power 6, so this is a second approximation. So, if the form of F is known, you can have the expressions for the order statistics and the amount of the error in this approximation is not much 1; some calculations have been done for various values of n and r for a specific distributions and then one can check that how much error is there.

And these approximations are quite satisfactory, especially if I consider the second order approximations. So, this approximation made use of the fact that Taylor expansion is possible and second thing is that we assume, of course that assumption is not a very stringent assumption because what we are having is that F function I am assuming to be strictly (()) (32:41), so that means basically I am taking to be a nice function.

And if that is so, I am assuming higher order moments to be negligible and in the sense that we are dividing by higher powers of n, so and the deviation will be much less because as you can see here itself in the denominators, we are getting the term like, n + 1 to the power 4, n + 2 square, so I take further, then this term will be very, very high and therefore, they become negligible, so it is not a stringent or you can say bad assumption here as such.

(Refer Slide Time: 33:22)



Next thing is that one can talk about the asymptotic distributions also, next we derive asymptotic distributions of Xr that is rth order statistics. Now, here there can be 2 cases; one is that only n turns to infinity, r remains fixed, so that means for example, I am finding out asymptotic distribution of the minimum, the maximum etc, so that means the position is fixed but in the second can be that the position can also vary.

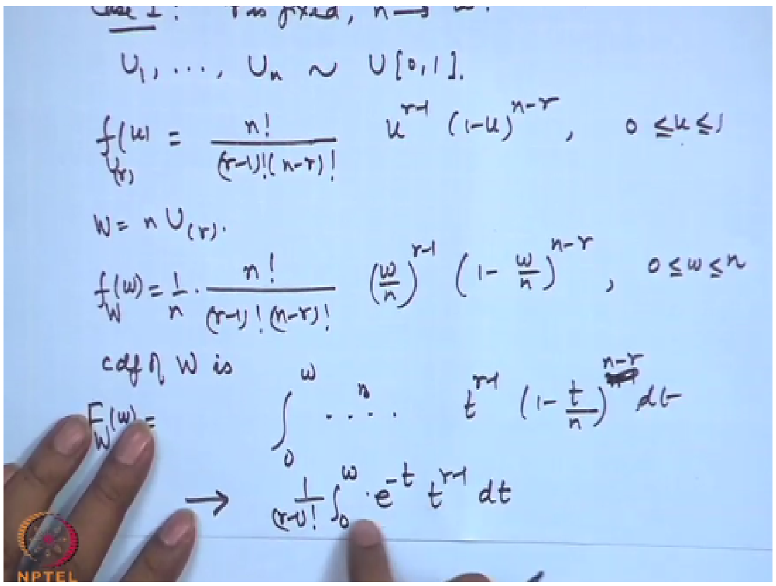
That means I take say r tends to infinity, n tends to infinity such that r/n tends to a fixed value, so I will consider these 2 cases; case 1, r is fixed and n is tending to infinity, so let us firstly consider the case of the order statistics from the uniform distribution, so I am having say, U1, U2, Un from uniform 0, 1. So, certainly here the distribution of the rth that was n factorial/ r-1 factorial n – r factorial u to the power r – 1, 1 – u to the power n –r, 0 < u <1.

Here, you take W to be say n times Ur, then what is the distribution of W, it will become n factorial/ r -1 factorial n - r factorial, then this is becoming w/ n to the power r -1, 1 – w/n to the power n - r and you will get 1/n extra term here, w is between 0 to n. Let us look at the cumulative distribution function of W that is F that is becoming equal to integral 0 to w and all these coefficients will becoming there.

You can actually consider say, say this is becoming basically n to the power r -1 factorial and this is w/ n to the power n – r, so if I integrate I will get; oaky let me write it here, so all these coefficients will be there, then I have t to the power r -1, 1 – t/n to the power r -; sorry n - r dt and other coefficients will be there, which I will not written here. If I take the limit of this as n tends to infinity, then what I get here?

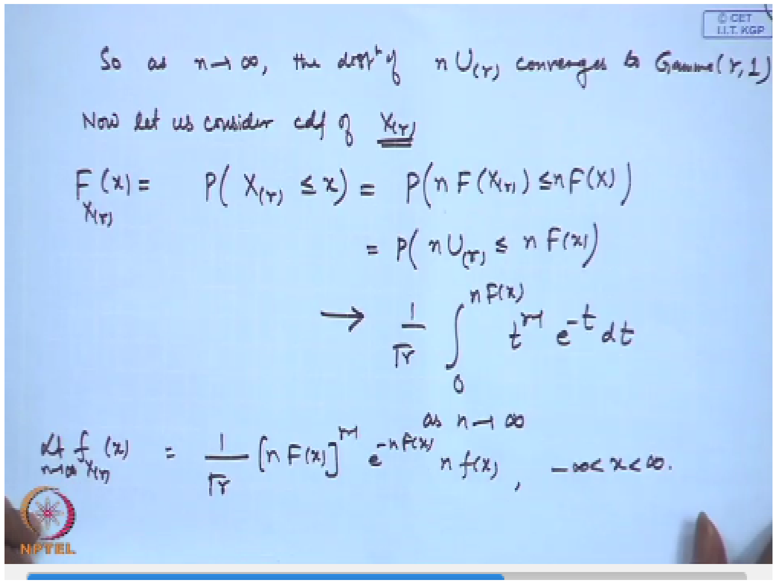
See this term will go to e to the power –t here and the first thing is that the limit will exist and this will go to e to the power -t because this time will go to 0 t/n tending to; so now if I take the limit of this coefficient here, you can check what are the terms that are coming here, you are actually getting n to the power r in the denominator and here you are having n * n - 1 * up to n – r + 1/ n to the; that means basically there are total r terms here.

(Refer Slide Time: 37:29)



In the denominator, you are getting n to the power r, so each of the terms will be adjusted like you have n/n, n – 1/n, n-2/n up to n – r + 1/n, each of these terms as n tends to infinity will converge to 1. So, if I take the limit here this is converting to exactly a term of the type 0 to w, 1/r - 1 factorial e to the power –t, t to the power r -1 dt, so this is very instructive, when you look at this, this is nothing but the CDF of a gamma distribution.

(Refer Slide Time: 37:53)



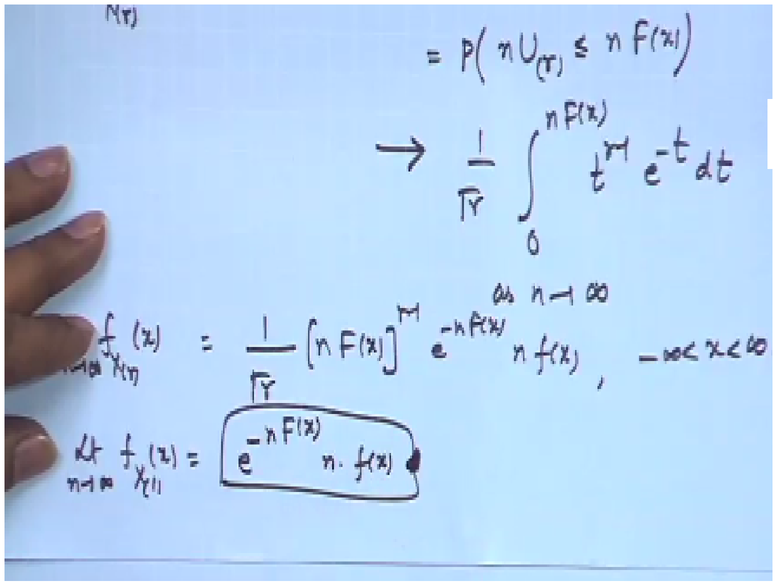
That means we are saying asymptotic distribution of; so as n tends to infinity, the distribution of n times Ur converges to gamma r1, a scale parameter 1 here, so that is one result that we are able to obtain. Now, let us consider CDF of Xr, okay. So, if you consider the CDF of Xr, what is happening here? F of Xr at some point say x, that is = probability of Xr <= x, which we can write as F of Xr <= Fx.

So, this =; you can multiple by n also, so this is nothing but n Ur <= n of Fx. Now, we are showing that this is converging to 1/ gamma r integral 0 to n Fx t to the power r -1 e to the power –t dt as n tends to infinity. So, this is very interesting actually, we are able to obtain the general form of the limiting CDF of rth order statistics as the sample size tends to infinity and if we can consider the derivative here, then limit of that also can be obtain as 1/gamma r n Fx to the power r -1 e to the power -n Fx nfx.

Because I am just applying the Leibniz rule here, so this is; the derivative of this becoming n fx and here n Fx to the power r -1 e to the power - n Fx, so we are able to obtain the general form of the limiting probability density function of the rth order statistics here. In the second part is; in this one, you can see that we have applied the condition that r is fixed but n is tending to infinity.

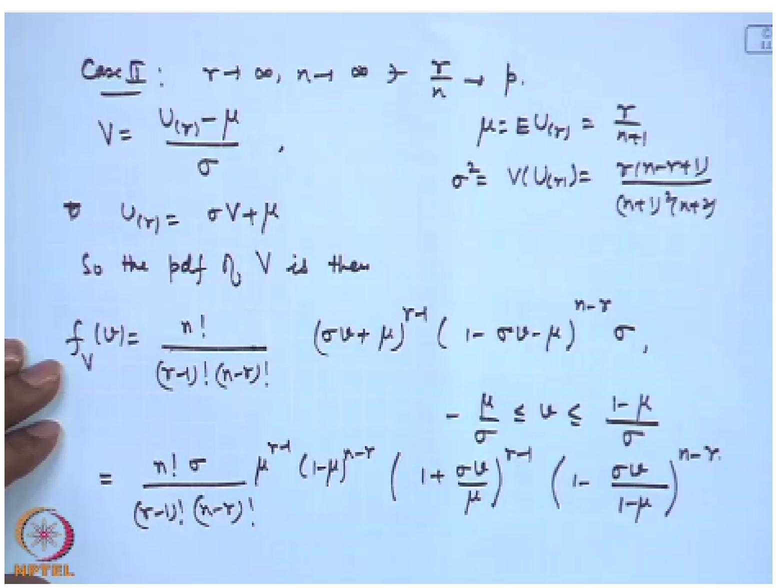
Now, the second can be that I can take both r and n to be tending to infinity such that r/n tends to a fixed number. Actually, what is the difference here? If r is fixed, that means I am finding say suppose for second one nor third one etc., but if I take r also tending to infinity that means I am fixing the position for example, it could be middle like median or a quantile, so there is a difference in the treatment here.

(Refer Slide Time: 41:20)



We may also see what happens to this thing, for example, if I look at say limiting density for the first one, then what happens? Then r is = 1 here, so this term will go away, this term will away, you get simply e to the power -n Fx n times fx, okay, so this is a very very interesting thing that asymptotic distribution for a general form for the minimum can be obtain, which is of course, you can say it is an exponential form.

(Refer Slide Time: 42:14)

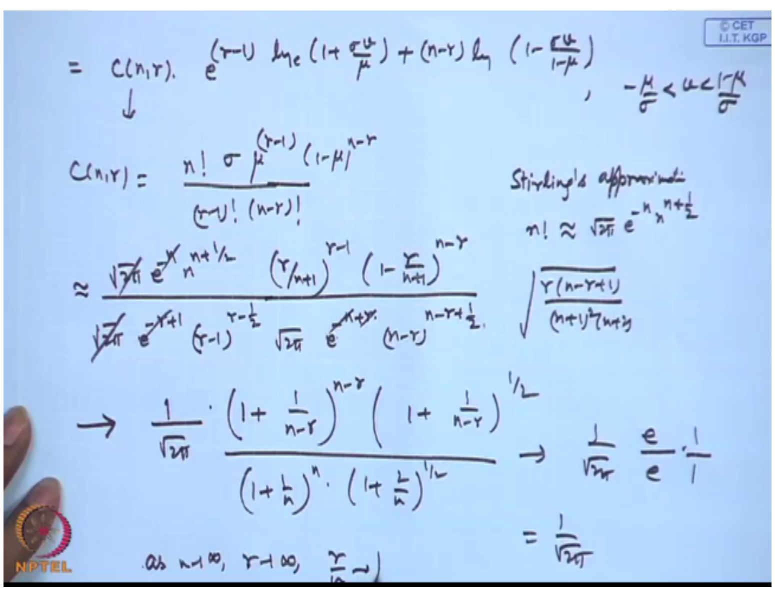


And that is why in the negative exponential distribution; basically you were getting n times that density was repeated there. Let us consider case 2; r tends to infinity, n tends to infinity such that r/n tends to p, okay. So, let us look at the Ur – Mu/ sigma, where Mu is the expectation of Ur that is r/n + 1 and sigma square is = variance of Ur that is = r * n – r +1/ n + 1 square * n + 2. So, the density of Ur is known, so let us write that.

Actually, I already considered this density here, so from here I can find out the density of Ur – Mu/ sigma because this will give me actually Ur is = sigma V + Mu, so the pdf of V is then fv that is = n factorial/r -1 factorial n - r factorial, then U is sigma v + Mu to the power r -1, 1 – sigma v - Mu to the power n - r and then you are also getting du is = sigma dv, another thing is the range here; u is between 0 to 1.

So, that will give me the range of v from –Mu/sigma to 1-Mu/ sigma that is –Mu/ sigma <= v < = 1- Mu/ sigma. So, let us write the terms in a slightly adjusted fashion, so n factorial sigma r -1 factorial n - r factorial, this Mu, I will take out from here. So, I get Mu to the power r -1, from here I take out 1 – Mu, then this term is becoming = 1+ sigma v/Mu to the power r -1 and 1- sigma v/ 1- Mu to the power n - r this.

(Refer Slide Time: 45:02)



This term I further expresses n factorial, so some coefficient is there; let me write it as a coefficient of n and r; e to the power r -1 log of 1+ sigma v/Mu + n – r times log of 1 - sigma v/1-

Mu and there are some coefficients here, which I not written here; Mu/ sigma < v < 1-Mu/sigma, this c, n, r coefficient is actually all these terms which are written there, so I am just omitting that here.

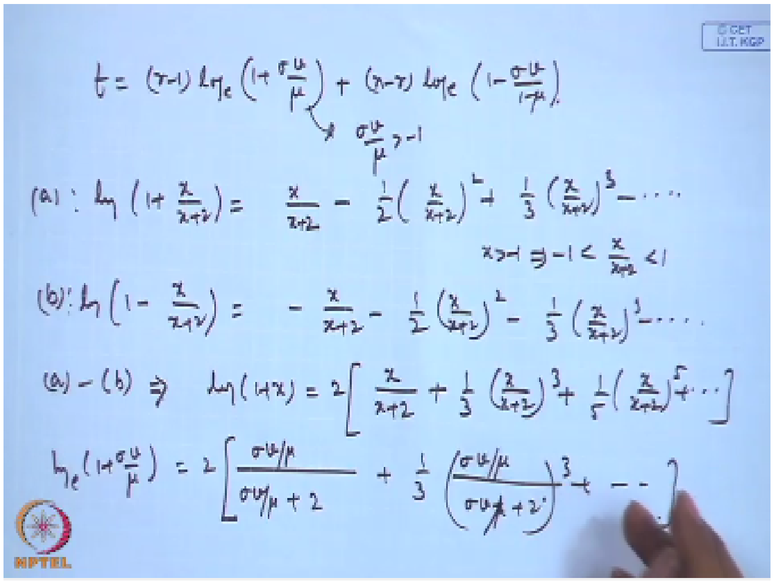
So, this c, n, r term, if you look at; this c, n, r, which I wrote as n factorial sigma Mu to the power r -1 * 1 – Mu to the power n – r/ r -1 factorial n – r factorial. If I take here are; r tending to infinity, n intending to infinity such that r/n tends to p, then this can be shown to be convergent to; so you can actually use some Stirling's approximation. If you remember the Stirling's approximation that is n factorial is approximated by root 2 pi e to the power –n, n to the power n + 1/2.

If we use this approximation, then this can be written as root 2 pi e to the power –n, n to the power n + 1/2/; since I am taking of r also tending to infinity, so another root 2 pi will come here eto the power -r + 1, r -1 to the power r- 1/2, then root 2 pi e to the power –n + r and then n –r to the power n – r + 1/2, then you have sigma, which is actually coming as the square root of r * n- r + 1/ n + 1 square * n + 2 and then you have Mu that is r/n + 1 to the power r-1, 1 – r/ n + 1 to the power n –r.

So, this cancels out e to the power - n cancels out, then e to the power –r and e to the power +r also cancels out, so you can actually show, I mean one can basically obtain the limit of this; this is actually converting to 1/ root 2pi, basically what you will get 1/root 2 pi term here and then all other terms can be adjusted easily1 + 1/n - r to the power n – r, then you have 1 + 1/n - r to the power 1/2 that is coming by division here.

Because see, I am getting n –r, n +1 – r here, so this I divide here, so I take the equal power which are coming that can be divided here and similarly in this one, so this I will get as 1 + 1/n to the power n and I will get 1 + 2/n to the power 1/2. So, now if I take the limit is simply 1/root 2 pi, then this will give e, in the denominator, I will get e, this will give me 1, this will give me 1, so this I will get a simply 1/root 2 pi.

(Refer Slide Time: 49:04)

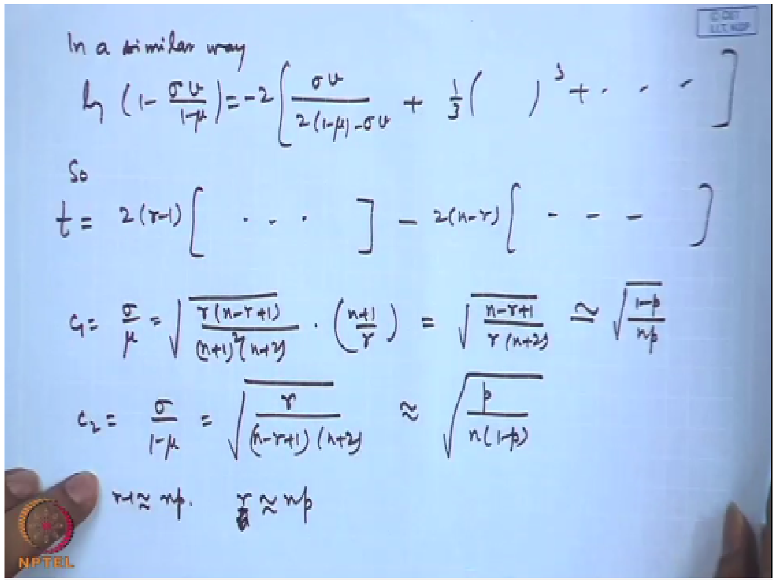


So, this coefficient c, n, r this converges to this as n tends to infinity, r tends to infinity such that r/n tends to p. of course r/n has not played a role here but because all the terms are not coming in the particular fashion. Now, let us look at this part. This part, if I consider as say, t that is r -11+ sigma v/ Mu + n - r log of 1 - sigma v/ 1 – Mu. If I consider this, so we can consider expansion of this type of forms.

Log of 1+x/x+2 that is = x/ x + 2 -1/2 x/ x + 2 square and so on; 1/3 x/ x +2 cube, this type of expansion will be valid, if I consider -1 < x/ x +2 < 1, that means x > -1. Similarly, I can consider log of 1 – x +; that is becoming – x/ x +2 -1/ 2 x/ x +2 square -1/3 x/ x + 2 cube and so on. So, if I take the difference that means let me call it, a and b, then if I do a – b, then I will get this as log of 1+ x = twice x/ x +2+1/3 x/ x + 2 cube 1 /5 x/ x + 2 to the power 5 etc.

So, for expansion of this, I am unable to use directly that this is between -1 and 1, therefore I am using another form here. This expansion will be valid provided x is > -1 but that is true here, here this sigma v/ Mu is > -1 and similarly this one will also be. So, if we consider this, then I get log of 1 + sigma v/ Mu is = 2 times sigma v/ Mu/; this is of course slightly cumbersome expression but is still one can write it in the closed form, so this cube and so on.

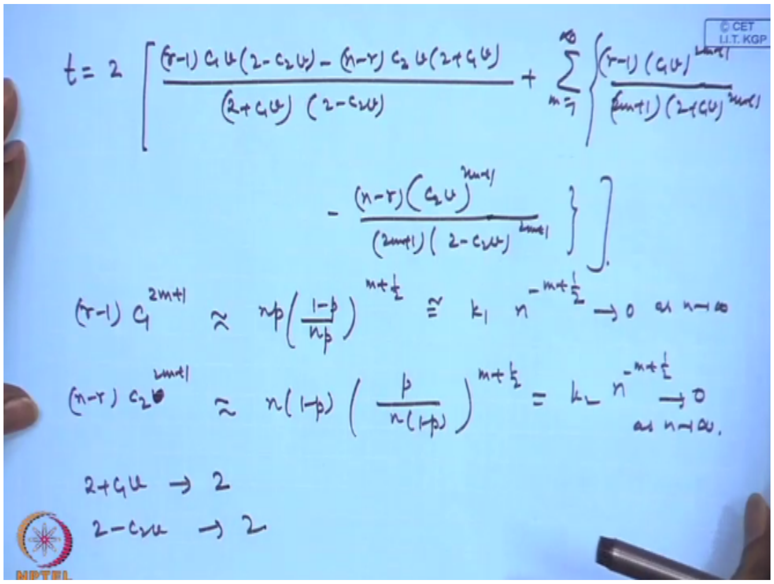
(Refer Slide Time: 51:52)



In a similar way, I can consider in a similar way, log of 1 - sigma v/ 1- Mu, so that is 2 times sigma v, with the – sign here, twice 1 – Mu - sigma v +1/3 cube and so on, all these type of terms will be coming. So, the term that I defined as t that is r -1 times this one + n -r times this, it will be a somewhat complicated expression but I can write it as 2 times r -1, this particular term - twice n - r times all these terms will come here.

Now, if I defined something like c1 is = sigma/ Mu that is = root r * n - r +1/ n + 1 square * n + 2, n + 1/ r, then this is actually n - r +1/ r * n + 2, then if I use r/n, is = approximately p, then it is equivalent to 1 – p/ np. In a similar way, I define other terms; sigma 2 that is sigma/ 1 - Mu that is root r/ n - r +1 * n + 2, so there is approximately p/ n * 1 - p as r tends to infinity, n tends to infinity, r/n tends to p.

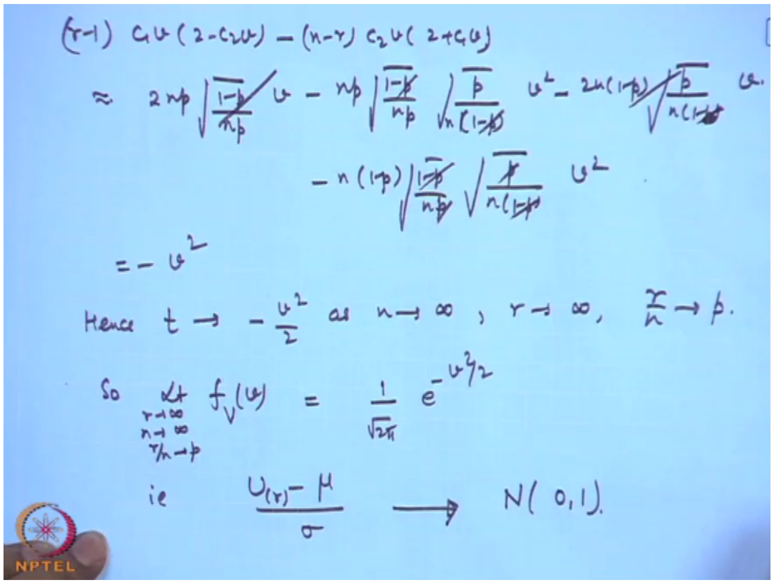
(Refer Slide Time: 54:19)



And in terms of this c1 and c2, of course r -1 is also equivalent to np, r/n is; r is equivalent to np etc, all these terms will be using, so let us substitute here. So, what I get t is; there are some more terms that will be appearing here, let me write it here, see this 2; t can be written as 2 times after simplification as r -1 c1 v, 2 - c2 v – n - r c2 v, 2 + c1 v/ 2+ c1 v * 2 - c2 v + summation m = 1 to infinity r -1 c1 v to the power 2m +1/ 2m+1 * 2+ c1 v to the power 2m +1.

This is one term and second term will be n - r c2 v to the power 2m + 1/ 2m +1 * 2 - c2 v to the power 2m +1. Now, this c1 and c2, I defined here and I obtained their asymptotic expressions here, there will be other terms coming there for example, r -1 c1 to the power 2m +1, then this will be approximately np/ 1- p/ np to the power m + 1/2, which we can of course write as say, k1 n to the power –n + 1/2.

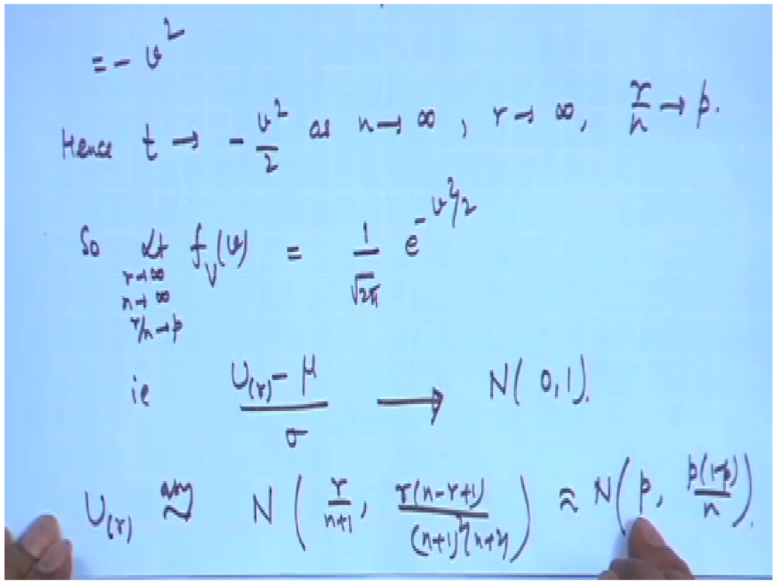
(Refer Slide Time: 56:32)



Of course, that will go to 0 as n tends to infinity, similarly n - r times c2 to the power 2m + 1 that is also approximately n times 1 – p, p/n * 1 - p to the power m + 1/2 that is again some k2 times n to the power –n + 1/2 that is again going to 0 as n tends to infinity. Similarly, if I consider say, 2 + c1 v that we will go to 2; 2 – c2 if I considered that will also go to 2, if I substitute all these values, then a rough expression for r -1 c1 v 2 – c2 v –n-r c2 v, 2 + c1 v.

That will be approximately twice np root 1 – p/ np v - np root 1 –p/np root p/ n * 1 - p v square -2n * 1 –p root p/ n * 1 –p v, –n * 1 – p root 1 –p/np, p/ n * 1 - p v square, so easily you can see this terms get cancelled out okay and this term will itself cancelled out with this term and we are left with actually –v square. So, what we are getting is that; t converges to -v square/2 as n tends to infinity, r tends to infinity and r/n tends to p.

(Refer Slide Time: 58:31)



So, what we proved is that if I considered the limiting distribution of; limiting pdf of the v; v was Ur – Mu/ sigma as r tending to infinity, n tending to infinity and r/n tending to p as 1/root 2 pi e to the power - v square/2, that is Ur – Mu/ sigma converges to normal 0, 1 that is a very, very significant result and we can say that asymptotic distribution of Ur that is rth order statistics is normal with mean Mu that is r/n +1 and variances sigma is square.

That is r * n - r+1/ n +1 square * n + 2, this is asymptotic distribution of this, which of course, you can say as p and this is r/n is approximately p and in the second part, you can write as 1 – p/n that is approximately p, p * 1 –p/n, so you can think of this as actually binomial type of term here, mean is np and the variance is np * 1 –p etc. Now, if I use a inverse function here, if I use a inverse function here, then this is becoming F inverse p.

That means it is the pth quantile coming here, so let me write it here it in the form of a theorem. So, basically, we are able to obtain the asymptotic distribution of Ur as the normal distribution here and therefore, if I use the inverse function which I will be giving in the tomorrow's; in the next lecture the asymptotic distribution of the rth order statistics under these assumptions is obtained as a normal distribution.

So, you have two cases; one is when r is tending to infinity, n is tending to infinity and r is fixed, in that case you are getting related to gamma function and here you are getting related to the normal distribution. I will be using this for obtaining confidence intervals for the population quantiles as I have mentioned that in the non-parametric situation rather than considering the means and the variances etc., we discuss the positional point that is the median, quartiles, the quantiles the percentiles etc. So, I will be doing it in the next class.