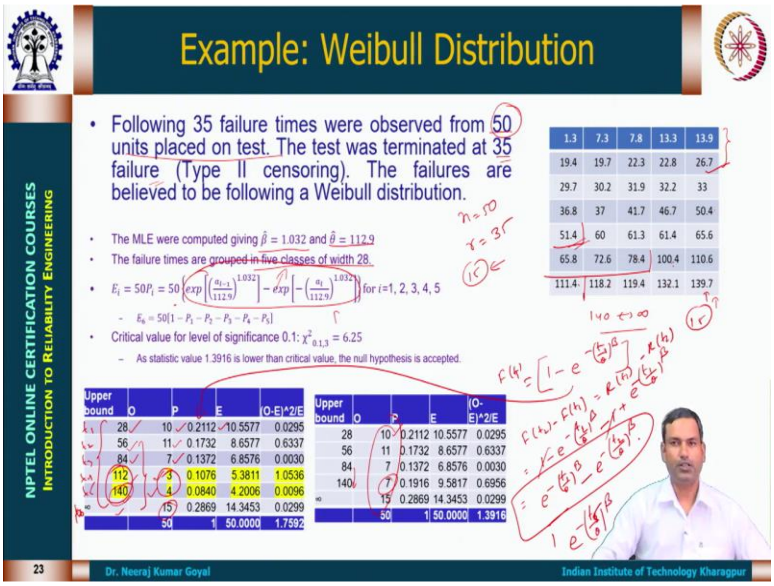
Introduction to reliability
Professor Neeraj Kumar Goyal
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur
Lecture-34
Goodness of Fit (GOF) Tests (Contd.)

(Refer Slide Time: 00:25)



Hello everyone. So, we have been discussing about Goodness of Fit Test in last class we discussed about the chi-square test. So, let us continue our discussion about chi-square test last time we have taken an example and we have seen that if we have the exponential distribution, how can we apply the chi-square test and see whether this fitting is acceptable or not.

Now, let us see the same thing for Weibull distribution. So, if let us assume that there are 35 failure times available and which has been put absorved by putting 50 units on tests. So, we have 50 units on test and we have 35 failure times.

So, that means, an 35th failure that it was sensored, whatever is the time at the 35th Failure same time will be applicable for rest of the 15 sensor devices. So, 15 devices are sensored, at the same point. So, here once we have this data we can use the MLE and MLE will give the value for beta and theta this we are as taking as a pre calculated value here.

So, once we have this beta and theta then now, we let us see that whether we are able to whether this is this fit to the Weibull is good or not. So, here my null hypothesis is that the failure time to fill failure the test fitting to Weibull distribution and alternative hypothesis is it is not following the Weibull distribution.

So, here failure times we have the 35 data we had. So, we have grouped in 5 classes. So, generally as we discussed earlier that if you divide the data into classes by Sturgis formula or some other. So, if you divided into 5 classes like like this, 1, 2 that is actually 6 classes. So, this is a 6 classes initially, if you use the Sturgis formula studies formula for 35 and 30 both comes around 6. So, we have the 6 classes.

Now, if we divide it into the 6 classes, we have in first class like 10 up to 28. So, if you see here up to 28 that means up to here, so, 5 plus 5, 10 failures are up to 28. Then second class is up to 56 so, 56 means here, that means, 5 plus 5,10 plus 1, 11, 11 failures are in up to 56.

Then for 84 that means 56 to 84 so, 84 is coming somewhere up to here that means, 5 plus 2, 7, 7 is lying into third interval and up to 112. So, for 112 is up to here 1, 2, 3 data points, 3 data points in 112 and up to 140 is because of a test was last failure was observed is 140. So, test continued up to 140, 35th failure was at 140. So, 139.7 almost same as the 140.

So, 112, 3, 4, 4 failures are in from 112 to 140 and hear the sensoring happened, that means the rest of the 15 failures were not happen, so that 15 devices were sensored at this point.

So, that means the 15 devices will fall somewhere from 150 to infinity, because these are not failed. So, these 15 devices will fail somewhere in 140 to 50. So, that comes out to be 15. As we see here, again, these 2 data points if we see 3 and 4, they are falling below the 5. So, what we have to do, we have to combine these 2. So, this interval, rather than taking from 84 to 112, and then 112 to 140, we will take 84 to 140 directly and so that number of failures will become 7 here.

Now, we have the 5 intervals here, we have the 5 classes. And in 5 classes, we have this number of failures, which are all above 5. So, now we can use the same formula as we did earlier. We find out the each class probability, probability of this class. How can we get, we can use this formula probability as we discussed earlier one minus e to the power minus t upon theta raised to the power beta.

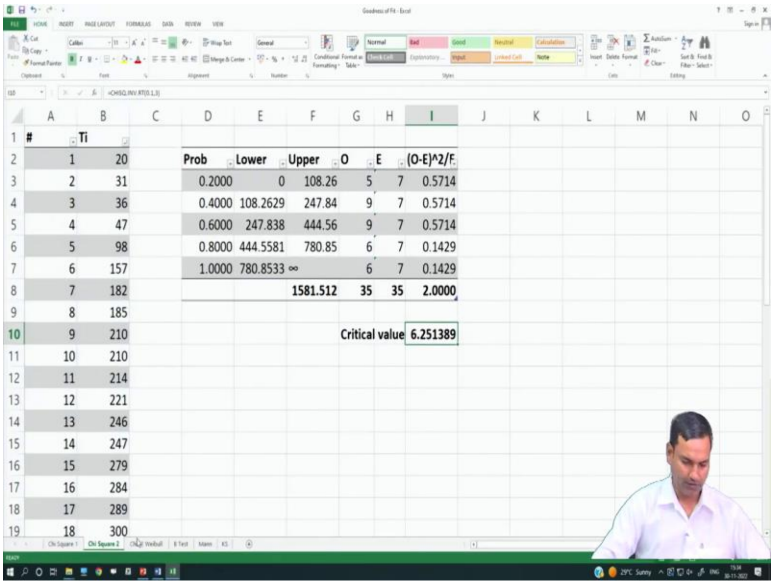
So, if I am if I want to know let us say this is t1 this is t2, this is t3, this is t4, this is t5, I want to know the failures between upto t1, so upto t 1 is F t1. So, this will be my first value this value would be coming here, the second interval, second interval is F t2 minus F t1 probability of falling in second interval is F t2 minus F t1 this is nothing but 1 minus e to the power minus t1, t2 upon theta raised to the power beta minus 1 plus e to the power minus t1 divided by theta raised to the power beta.

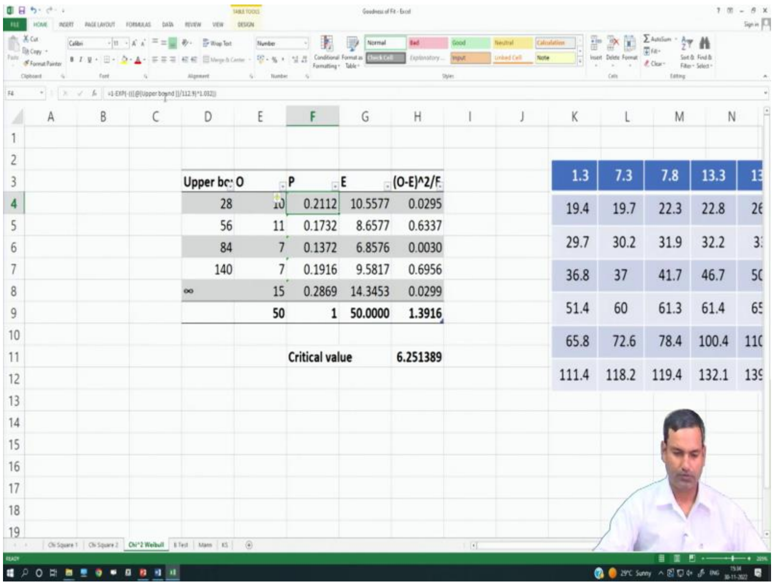
So, 1 and 1 can get cancelled. So, in a way we can write it as R t1 minus R t2 so, that means, e to the power minus t1 upon theta raised to the power beta this is plus sign so, this will come first and this is negative sign this come next, minus t2 divided by theta is to the power beta.

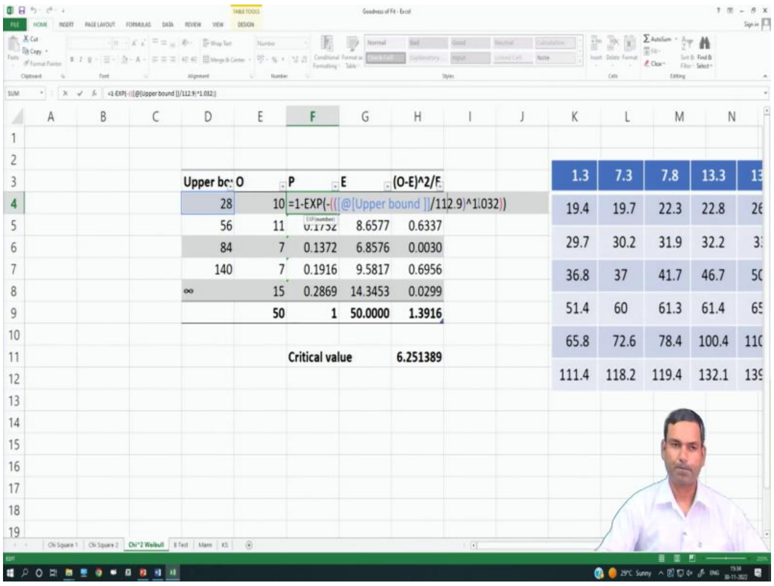
So, this we can apply a formula here and this formula will give me the probability of these interval intermediate interval and last interval is nothing but one last interval is the reliability that means, all the failures which are not happening up to 140 that will go to the last region. So, that will be nothing but e to the power minus this my t6 if I say, that is t6 minus t6 divided by, sorry, not t6, t5, t5 divided by theta raised to the power beta.

So, these calculations I have made already which is shown here, I will show the same thing in Excel because I have copied this all from Excel which I have already solved here.

(Refer Slide Time: 06:55)

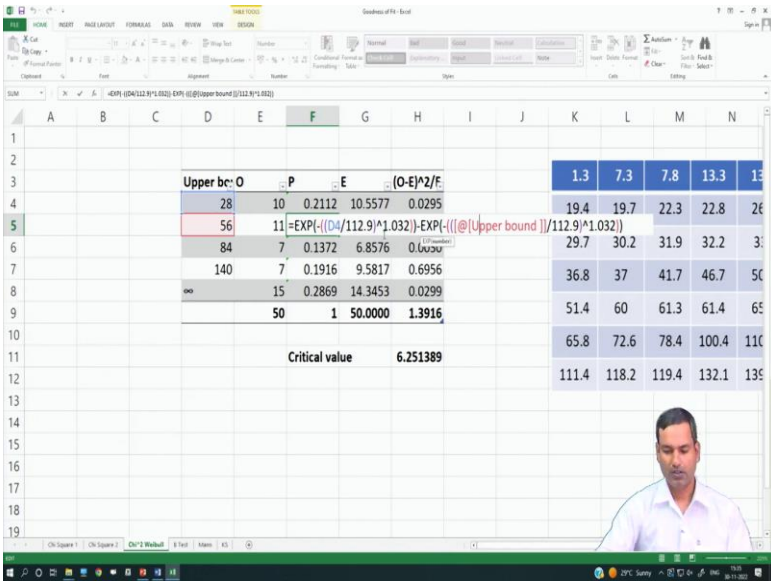


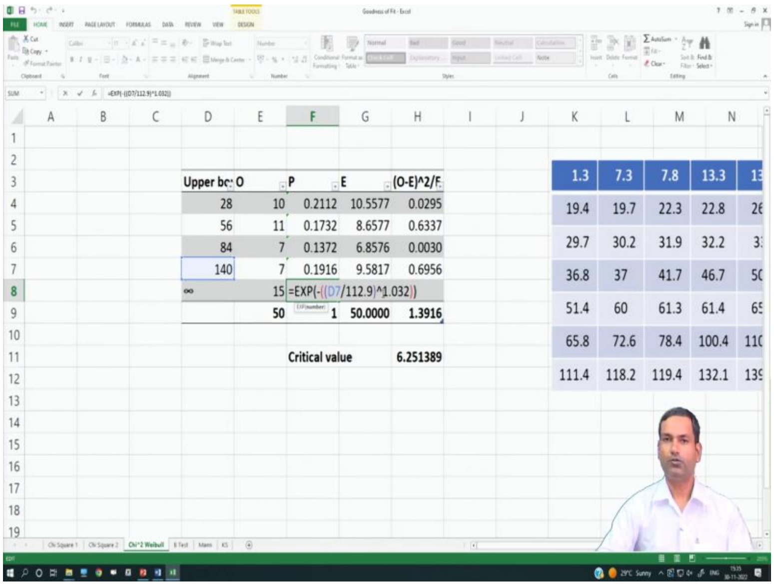


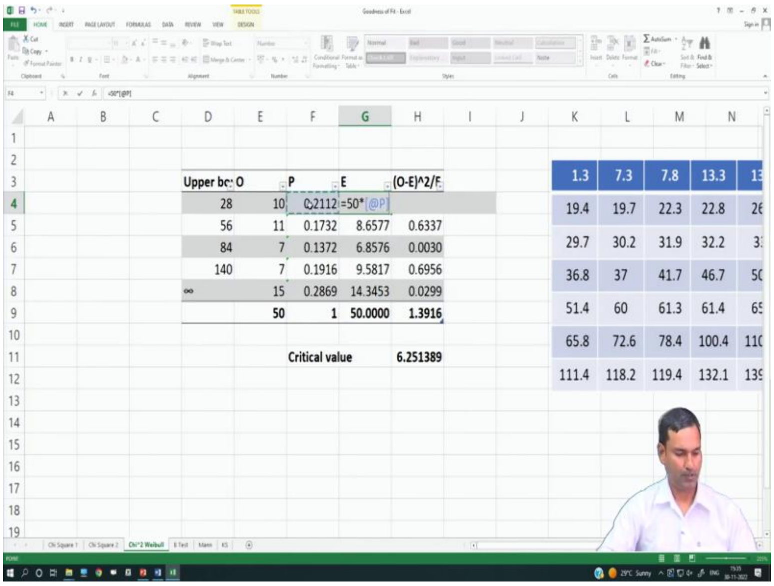


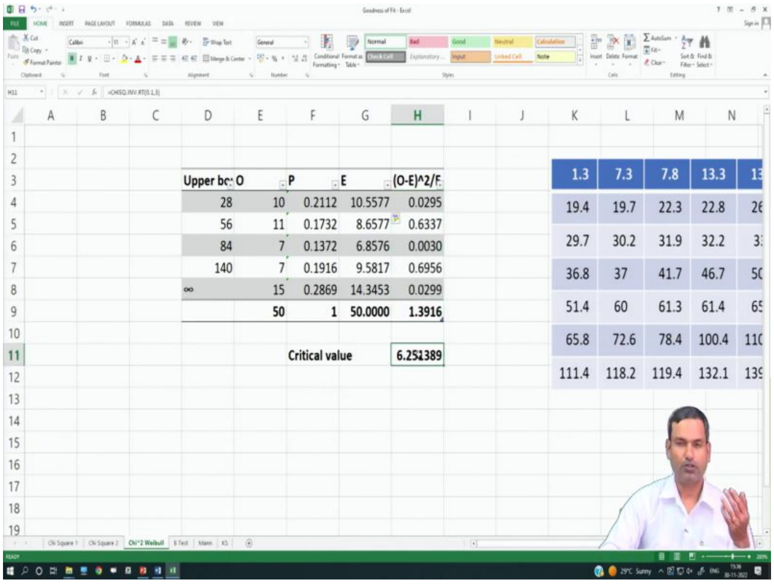
So, here if I see chi-square Weibull here so, we had this, n number of observations also we already seen. Now, let us see if I, how did we calculate the P, for first P that is 1 minus that is equal to n lambda and this is 11 see here theta is 112.9 and beta is 1.032. So, 1 minus e to the power minus t divided by theta raised to the power whole betta that is my first region.

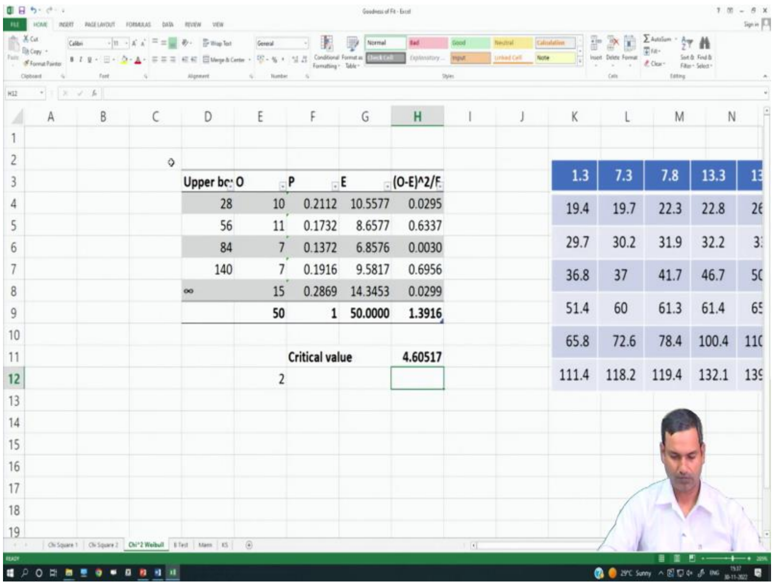
(Refer Slide Time: 07:33)






Upper bo O P E (O-E)^2/E
28 10 0.2112 =50*[@P]
56 11 0.1732 8.6577 0.6337
84 7 0.1372 6.8576 0.0030
140 7 0.1916 9.5817 0.6956
∞ 15 0.2869 14.3453 0.0299
50 1 50.0000 1.3916
Critical value 6.251389
1.3 7.3 7.8 13.3
19.4 19.7 22.3 22.8
29.7 30.2 31.9 32.2
36.8 37 41.7 46.7
51.4 60 61.3 61.4
65.8 72.6 78.4 100.4
111.4 118.2 119.4 132.1


Upper bo O P E (O-E)^2/E
28 10 0.2112 10.5577 0.0295
56 11 0.1732 8.6577 0.6337
84 7 0.1372 6.8576 0.0030
140 7 0.1916 9.5817 0.6956
∞ 15 0.2869 14.3453 0.0299
50 1 50.0000 1.3916
Critical value 6.251389
1.3 7.3 7.8 13.3
19.4 19.7 22.3 22.8
29.7 30.2 31.9 32.2
36.8 37 41.7 46.7
51.4 60 61.3 61.4
65.8 72.6 78.4 100.4
111.4 118.2 119.4 132.1


Upper bo O P E (O-E)^2/E
28 10 0.2112 10.5577 0.0295
56 11 0.1732 8.6577 0.6337
84 7 0.1372 6.8576 0.0030
140 7 0.1916 9.5817 0.6956
∞ 15 0.2869 14.3453 0.0299
50 1 50.0000 1.3916
Critical value 4.60517
2
1.3 7.3 7.8 13.3
19.4 19.7 22.3 22.8
29.7 30.2 31.9 32.2
36.8 37 41.7 46.7
51.4 60 61.3 61.4
65.8 72.6 78.4 100.4
111.4 118.2 119.4 132.1

And second region as you see that is exponential for t 1 minus that is lattice t 1 minus lattice t 2 that exponential for t 1 is D4 and t 2 is D5 divided by same 112.92 is to the point 1 minus r raised to the power 1.032. So, this gives me the second same will be followed here and third one will be nothing but the reliability at 140 that will be exponential minus D7 divided by this.

So, we are able to get the probability for each interval, this interval probability summation will always be 1 and how do I get the expected number of failures expected number of failures is nothing but the number of devices which I have put on test number of devices which I have put on tests for 50.

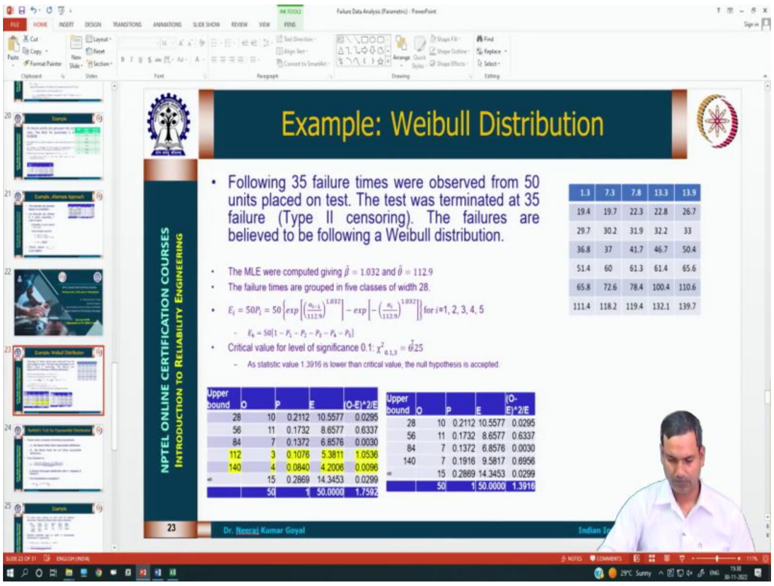
So, this will be equal to 50 multiplied by probability. And this gives me the number of fail expected number of failures because of this distribution fitting from the distribution fitting how much failures I am expecting in each failed, each interval here. And then again, I will use the same formula for chi-squared statistic that is O minus E whole squared divided by E these values comes out to be here and once I take a summation of this, this becomes 1.3916.

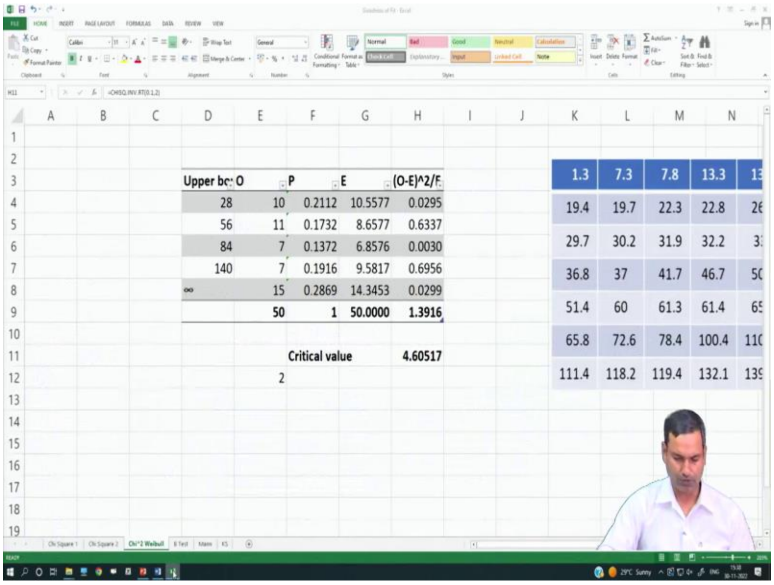
So, if I want to know the critical value critical value, I will take the number of intervals is 1, 2, 3, 4, 5 from 5, 1 is lost, because, 1 because 1 interval is already known, if other intervals are known, like if you know if I know the 4 interval fifth interval is already known, how much failures will be there in the fifth interval.

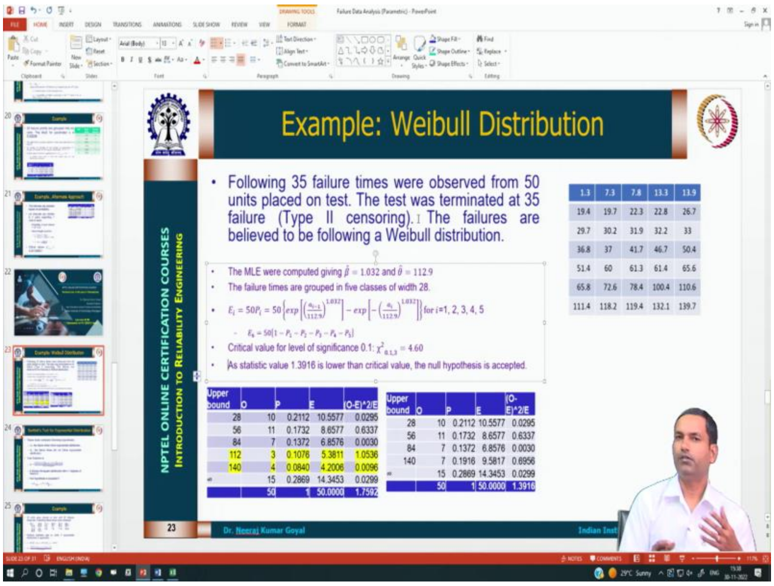
So, one degrees lost there and other degree is lost in estimation of parameters Weibull distribution is 2 parameters. So, from 5 this will be equal to 5 minus 1 for the grouping part and minus 2 for the number of parameters.

So, I have critical value so, 1, 2, 3, 4, 5, yes, so, my number of, my this what I say number of degrees of freedom becomes 2. So, this will be critical value will be 0.12. This was because it was a Weibull distribution. So, here critical value become point 4.6 which is much lower much higher than the 1.39, since our statistic value is lower than the critical value our null hypothesis is accepted and because of a null hypothesis accepted we can say that this data follows the Weibull distribution.

(Refer Slide Time: 10:32)

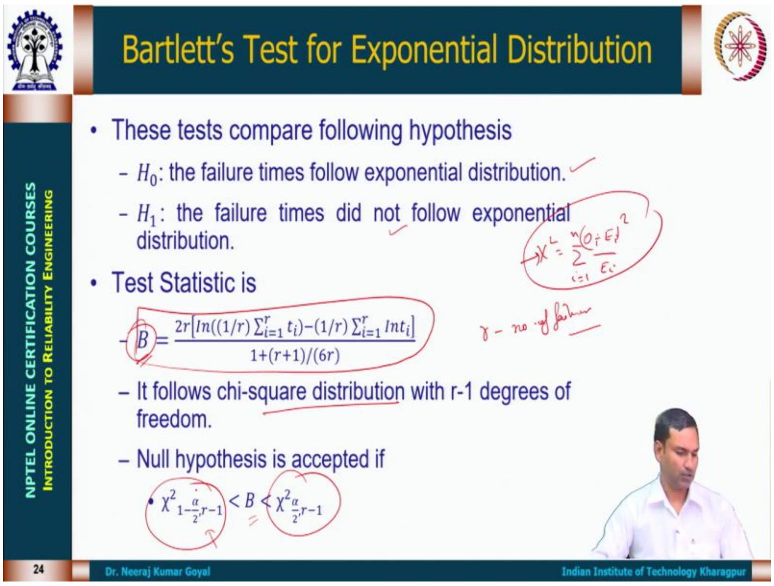






So, same thing we have got here these values were wrong. So, that guys, which I took 6.25 rather than 6.25 should be 4.60. So, here again as we have seen that chi-square test, same as we have done for normal, we have done for exponential Weibull you can do it for any distribution, same process, because it is not dependent on the distribution, you just need to divide and see in each interval, how many failures are the expected and how many failures are actually happening, but as we see here, for chi-square distribution, we require large data size, if sample size is small, we will not be able to do this properly.

(Refer Slide Time: 11:29)



So, for that purpose for different different distributions some specific test are suggested. That for exponential distribution Bartlett's test is used. So, by using the Bartlett's test, you can be you can find out whether the data is following the exponential distribution or it is not following the exponential distribution.

So, for Bartlett’s test like for chi-square test, this was a specific distribution which could be uneven, but here we are talking about the exponential distribution only. So, the null hypothesis is that the failure time follows the exponential distribution and alternate hypothesis is that times failure times does not, do not follow the exponential distribution.

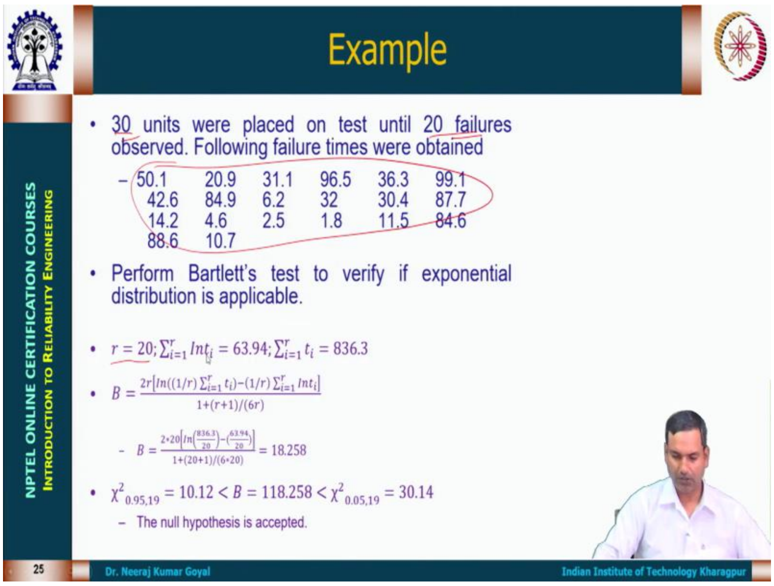
Like chi-square test we had the chi squared value was equal to O minus E, O i minus E i whole squared divided by E i and summation over i equal to 1 to n, this was our statistic, here the statistic is B for Bartlett. So, Bartlett test has this statistic where, this is like here we have the data. So, generally r is the number of failure. And so, if you do not have here it this is also applicable for sensor data.

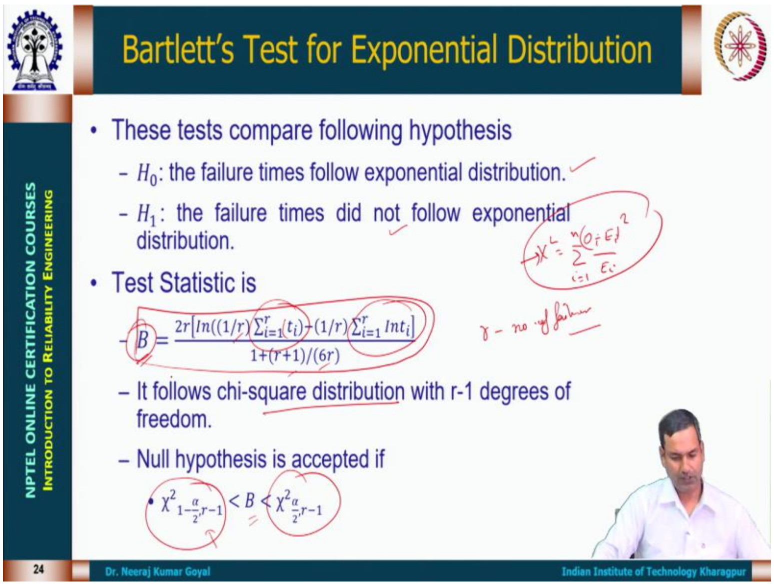
So, r is the failures data only the time number of failures, you may not have all the failures, then also you can use the same. So, it follows the chi square, this statistic which you have used this B also follows the chi-square distribution. So, the distribution would be same, but the statistic is different the calculation which you are making this statistic chi-square is also following chi-square distribution, this statistic B also follows the chi-square distribution.

So, the critical value we will obtain from the chi-square distribution and how the Chi but here the null hypothesis will be only accepted by the 2 sided comparison that means, the B value should be higher than the left limit and lower than the right limit. So, left limit is 1 minus alpha by 2 r minus 1 and right limit is alpha by 2 r minus 1.

So, and that is also again right side value that is leaving the area on the right hand side. So, this is the lower limit, lower bound this is the upper bound. So, if the, if my B values falling within this region I will accept that my time to failure are following the exponential distribution.

(Refer Slide Time: 14:10)

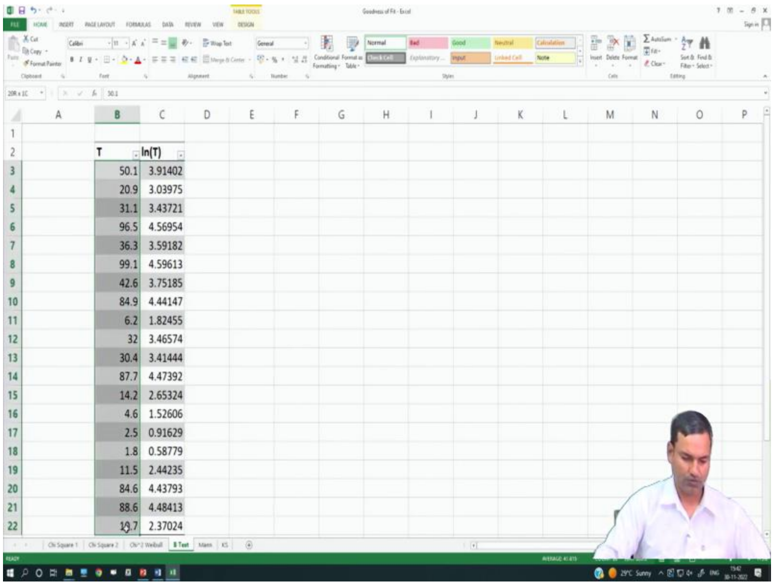


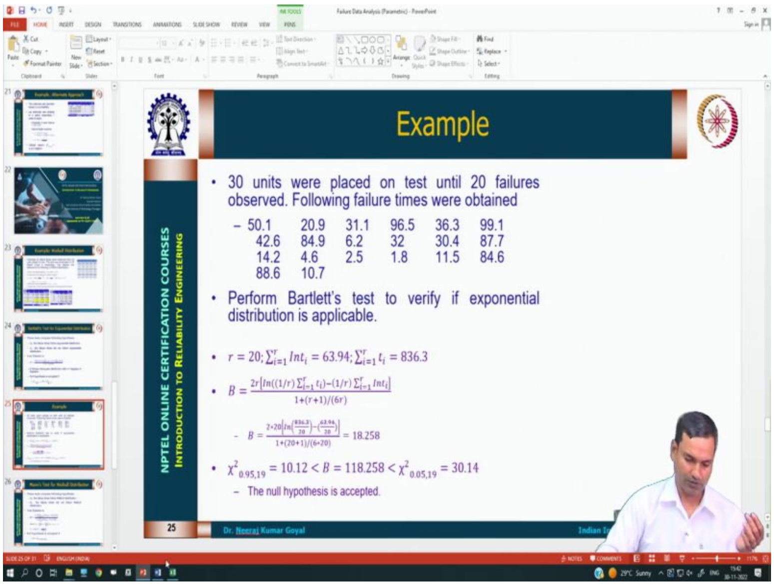


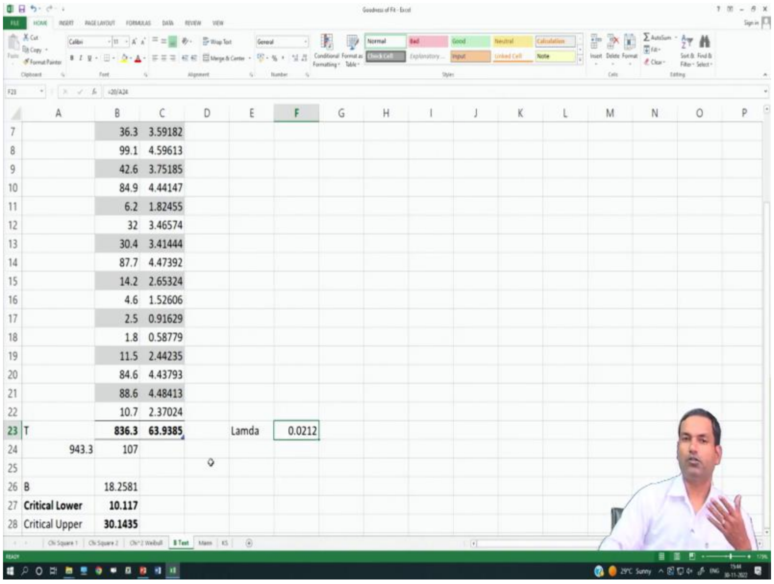
Otherwise, we may assume that so, let us take an example that 30 units were put on test out of which 20 failures happened. So, 10 units were sensored, but by just looking at the time to fill the data also we can find out whether this is following the exponential distribution or not. So, we want to perform the Bartlett test.

So, Bartlett's test if you want to do, how much is r number of failures that is 20. So, and as we see here for Bartlett test, we want to sum up 2 things, one is t i submission and other is ln of t i submission for number of failures. And we need r value rest of all are, so, if we know the r value submission of t i submission of ln of t i, we will be able to calculate the B value.

(Refer Slide Time: 15:04)







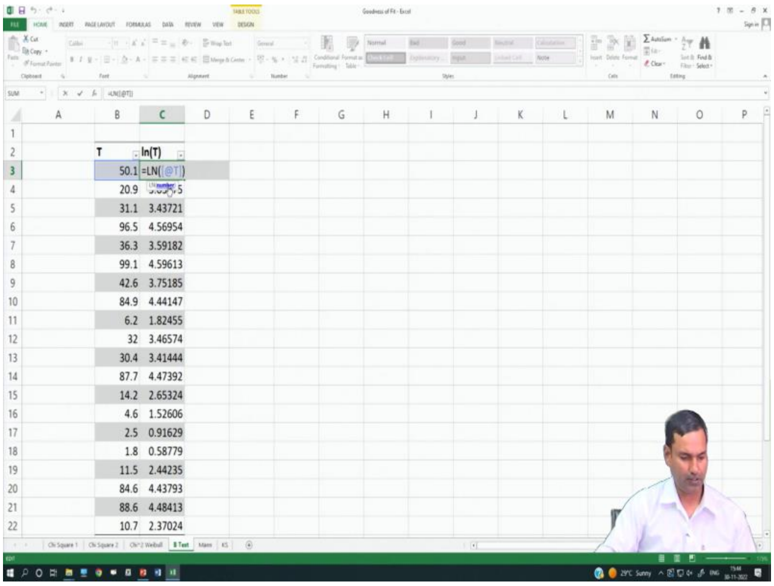
So, to do the same this is given here but I have done this exercise in Excel sheet also like this is my data T. So, these are my 20 failures if I want to know the mean, mean would be 836. So, not required right now, but we can do that if required let us say for this data since it though it is not part of problem, let us see, I want to know what is my because this is the exponential distribution I am assuming how much is the lambda if I want to calculate for this.

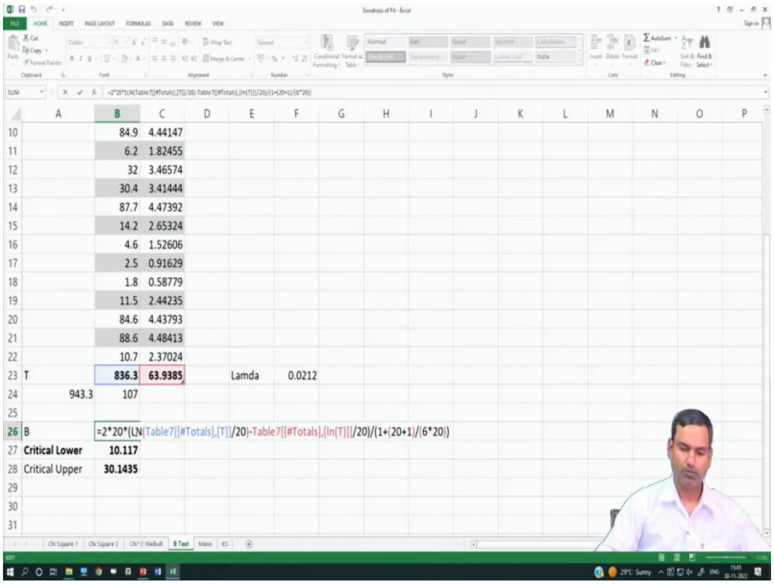
So, if I want to calculate the lambda then how do I calculate the lambda first I take the summation of failure time and then I also take the summation of sensor time, let us assume that all the failures would, whenever last failure was observed at that time the test was terminated that means, rest of the 10 devices only work for this much 10.7 hours. So, this is cumulated 107.

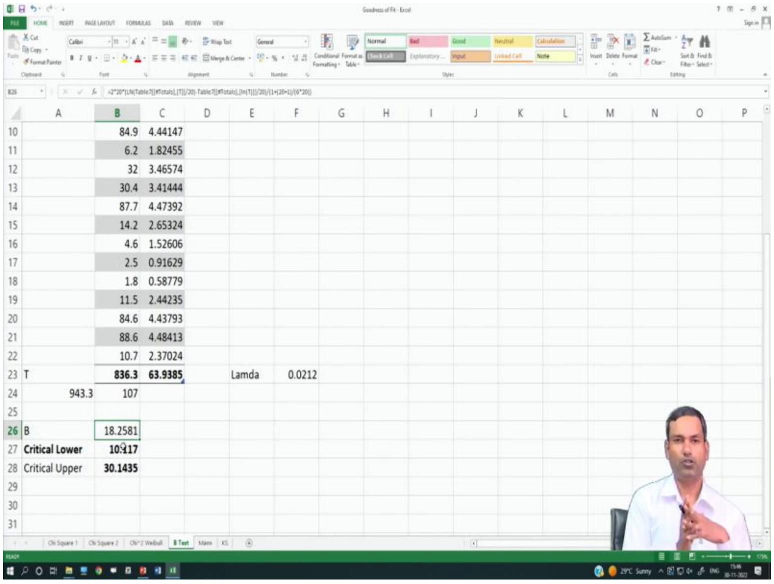
So, how much is total time total time is equal to 836 plus 107 943.3 and how many failures I observed. So, this becomes time and how much is r, r is 20. So, my lambda, lambda is equal to r divided by so, r is 20 and total time cumulative time is 943 though failure time submission is only 836 but there is a 10 devices which have worked for 10.7 hours each without failure. So, that time also need to be summed up added up with this. So, total time becomes 943.3 and my lambda comes out to be 0.0212.

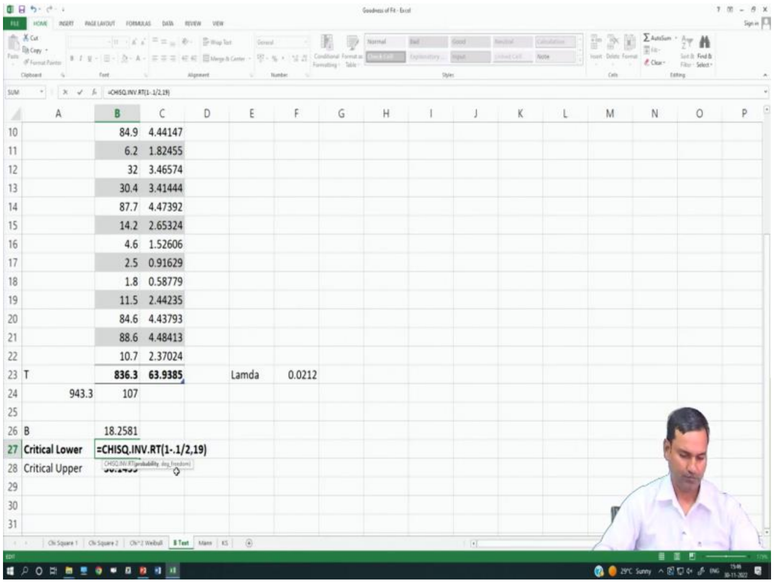
So, this is now how we use MLE for calculating the lambda, when sensor devices are there, so, this is just an example, we want to just check that whether this data is following the exponential distribution or not.

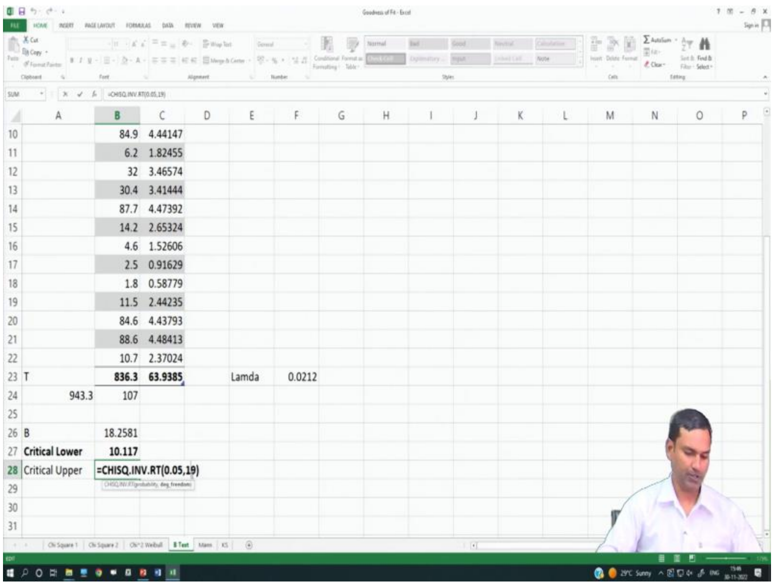
(Refer Slide Time: 17:22)

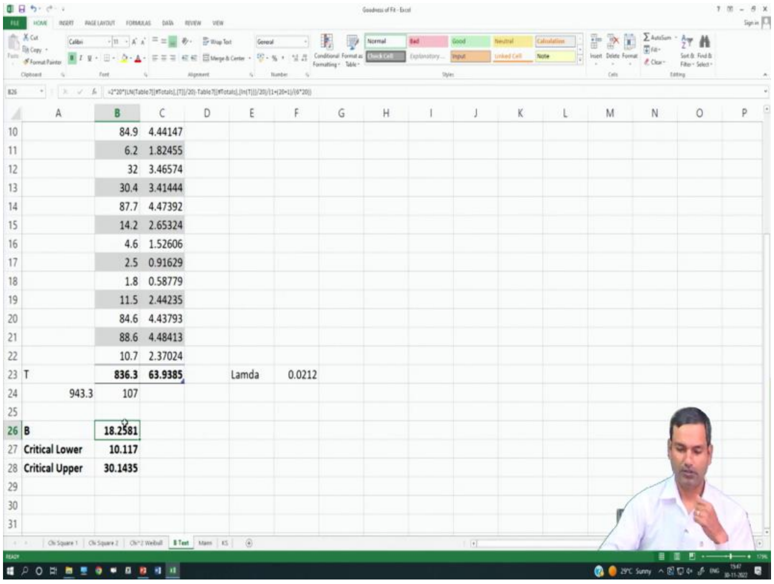





Goodness of Fit - Excel
10 84.9 4.44147
11 6.2 1.82455
12 32 3.46574
13 30.4 3.41444
14 87.7 4.47392
15 14.2 2.65324
16 4.6 1.52606
17 2.5 0.91629
18 1.8 0.58779
19 11.5 2.44235
20 84.6 4.43793
21 88.6 4.48413
22 10.7 2.37024
23 T 836.3 63.9385 Lamda 0.0212
24 943.3 107
26 B 18.2581
27 Critical Lower 10.117
28 Critical Upper 30.1435


Goodness of Fit - Excel
10 84.9 4.44147
11 6.2 1.82455
12 32 3.46574
13 30.4 3.41444
14 87.7 4.47392
15 14.2 2.65324
16 4.6 1.52606
17 2.5 0.91629
18 1.8 0.58779
19 11.5 2.44235
20 84.6 4.43793
21 88.6 4.48413
22 10.7 2.37024
23 T 836.3 63.9385 Lamda 0.0212
24 943.3 107
26 B 18.2581
27 Critical Lower =CHISQ.INV.RT(1-.1/2,19)
28 Critical Upper


Goodness of Fit - Excel
10 84.9 4.44147
11 6.2 1.82455
12 32 3.46574
13 30.4 3.41444
14 87.7 4.47392
15 14.2 2.65324
16 4.6 1.52606
17 2.5 0.91629
18 1.8 0.58779
19 11.5 2.44235
20 84.6 4.43793
21 88.6 4.48413
22 10.7 2.37024
23 T 836.3 63.9385 Lamda 0.0212
24 943.3 107
26 B 18.2581
27 Critical Lower 10.117
28 Critical Upper =CHISQ.INV.RT(0.05,19)



So, to do that, we need the 2 values, we would need the summation of T i and we need the summation of log, log of T i. So, T i is already here we will calculate log of I here that is ln of T i ln of T i and if we take the both summation this becomes ln of summation of T i this becomes summation of ln of T i.

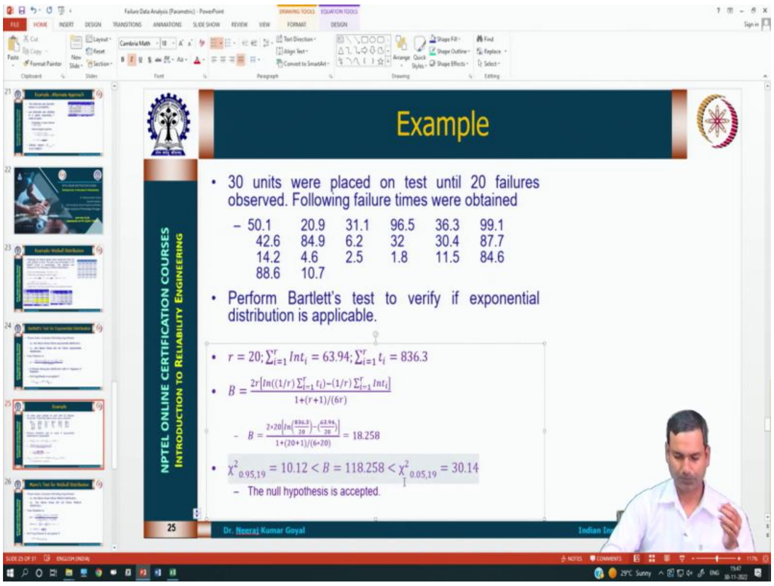
So, my B value which I have used here, that is formula is 2 into r, r is 20 here multiply by ln of summation of T divided by 20 minus, minus ln of t summation of ln of t divided by 20. Two things one is summation of T divided by 20 and another is summation of ln of I am taking log of that so, log of summation of T divided by 20 minus summation of log of values divided by 20 and whole divided by 1 plus r plus 1 divided by r plus 1 is only divided by 6 into r. So, 21 divided by 6 into 20 this if we slove my value of B comes out to be 18.

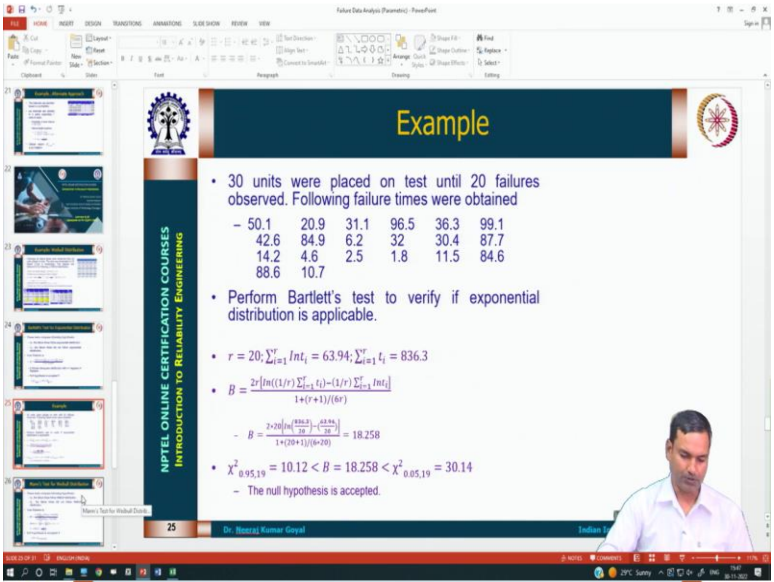
Now, whether my value, B value is falling in the region or not. So, for that we have to do the chi-square assessment. So, for chi-square we have the r minus 1 degrees of freedom here because only the number of intervals is there. So, here only one value is lost.

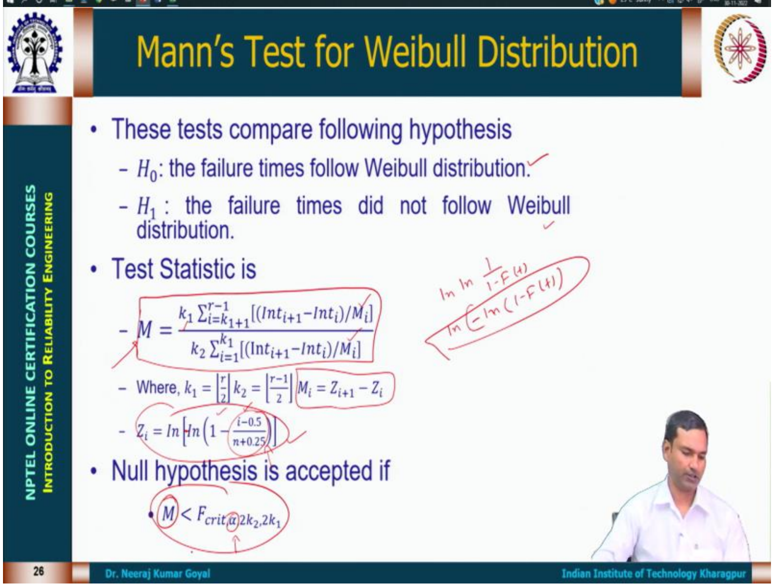
So, we have so, we want to know 1 minus alpha by 2 alpha is 10 percent 0.1 so, 1 minus 0.1 divided by 2 is 0.95, if I if you want I can write it here is 1 minus 0.1 divided by 2, 1 minus alpha by 2 you will get the same and right side value is alpha by 2 that is 0.1 divided by 2, 0.05 and degrees of freedom is 19.

Now, if you see here this might B value is well within this region. So, that means, my null hypothesis is accepted or I can say my data is following the exponential distribution. So, I can assume that my data is following the exponential distribution.

(Refer Slide Time: 19:58)







Same thing what I have done here? Same thing is given here this is not 118, this is 18. So, null hypothesis is accepted. Similarly, we may have another test for Weibull distribution. So, these are specific distribution like we use them Bartlett for exponential, we can use Mann's test for the Weibull distribution.

So, we can find out whether data is following the Weibull distribution or not. So, our null hypothesis is times to failure shows the follows the Weibull distribution and alternate is that time to follow does not time to fill it does not follow the Weibull distribution.

What is the statistic here, the statistic given by Mann's is this M equal to this and what is k 1 here k 1 is floor value of r by 2 that means, whenever I divide r by 2, I have to take only integer part the everything coming after the decimal point is dropped out and k 2 is r minus 1 by 2 that means subtracting 1 then dividing by 2.

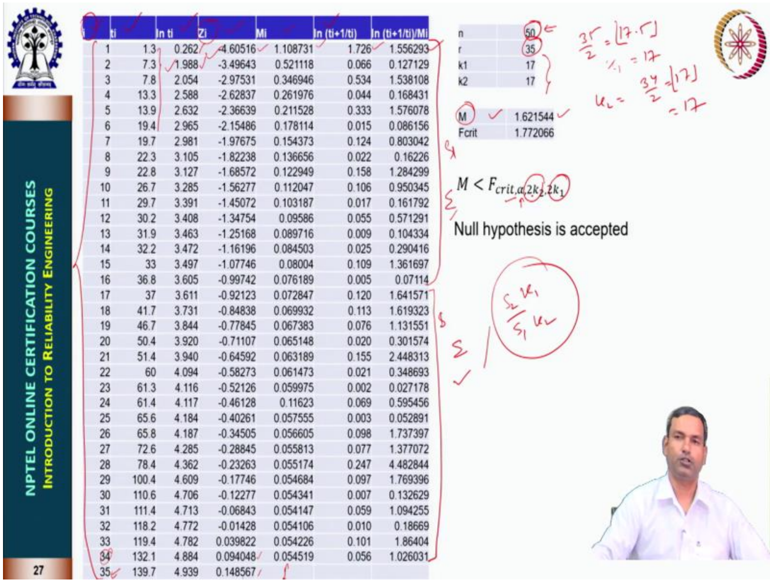
So, definitely if it is even number then you will get the integer if it is odd number then point 15 will be dropped and you will have the number. So, you will calculate, what is Mi here like we are using Mi here, Mi here this Mi is nothing but Zi plus 1 minus Zi and what Zi, Zi is like as we have seen earlier that is like we saw for the Weibull distribution ln of ln 1 upon 1 minus Ft. So, here so this is minus ln of minus ln this minus sign is missing here I will add that ln of minus ln or I can say ln of minus ln 1 minus Ft.

So, this value will give me and Ft is how much this, here we were using I minus 0.2 divided by n plus 0.4 if you want you can use that this is another statistic which is you would like if you change the value of n you will get, you will get this another set of the way of doing the median ranking.

So, here 0.5 and 0.25 is used. So, since he suggested 0.5, 0.25 we can go with that otherwise, you can use 0.3 and 0.4 also i minus 0.3 divided by n plus 0.4. So here this gives us the Zi. So, which is very similar to what we have done in the LSC and this value here this M follows the considered to be following the F distribution.

So, critical value is supposed is coming from the F distribution here. So, F distribution critical value, we have to take the right side distribution for alpha, alpha is my significance level 0.1 or 0.05. So, if the value of M falls below this it is accepted if it is not falling, if it is equal or higher than that, then it is not accepted.

(Refer Slide Time: 23:09)



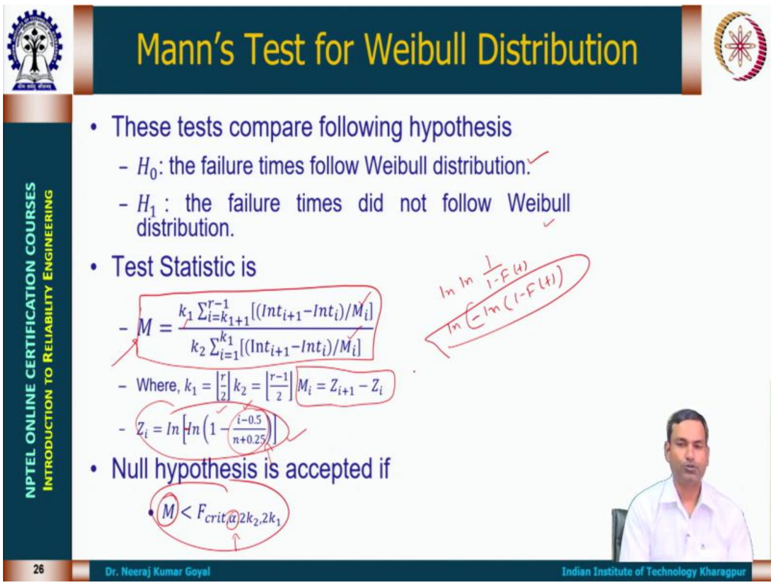
i	ti	ln ti	Zi	Mi	ln (ti+1/ti)	ln (ti+1/ti)/Mi
1	1.3	0.262	-4.60516	1.108731	1.726	1.556293
2	7.3	1.988	-3.49643	0.521118	0.066	0.127129
3	7.8	2.054	-2.97531	0.346946	0.534	1.538108
4	13.3	2.588	-2.62837	0.261976	0.044	0.168431
5	13.9	2.632	-2.36639	0.211528	0.333	1.576078
6	19.4	2.965	-2.15486	0.178114	0.015	0.086156
7	19.7	2.981	-1.97675	0.154373	0.124	0.803042
8	22.3	3.105	-1.82238	0.136656	0.022	0.16226
9	22.8	3.127	-1.68572	0.122949	0.158	1.284299
10	26.7	3.285	-1.56277	0.112047	0.106	0.950345
11	29.7	3.391	-1.45072	0.103187	0.017	0.161792
12	30.2	3.408	-1.34754	0.09586	0.055	0.571291
13	31.9	3.463	-1.25168	0.089716	0.009	0.104334
14	32.2	3.472	-1.16196	0.084503	0.025	0.290416
15	33	3.497	-1.07746	0.08004	0.109	1.361697
16	36.8	3.605	-0.99742	0.076189	0.005	0.07114
17	37	3.611	-0.92123	0.072847	0.120	1.641571
18	41.7	3.731	-0.84838	0.069932	0.113	1.619323
19	46.7	3.844	-0.77845	0.067383	0.076	1.131551
20	50.4	3.920	-0.71107	0.065148	0.020	0.301574
21	51.4	3.940	-0.64592	0.063189	0.155	2.448313
22	60	4.094	-0.58273	0.061473	0.021	0.348693
23	61.3	4.116	-0.52126	0.059975	0.002	0.027178
24	61.4	4.117	-0.46128	0.11623	0.069	0.595456
25	65.6	4.184	-0.40261	0.057555	0.003	0.052891
26	65.8	4.187	-0.34505	0.056605	0.098	1.737397
27	72.6	4.285	-0.28845	0.055813	0.077	1.377072
28	78.4	4.362	-0.23263	0.055174	0.247	4.482844
29	100.4	4.609	-0.17746	0.054684	0.097	1.769396
30	110.6	4.706	-0.12277	0.054341	0.007	0.132629
31	111.4	4.713	-0.06843	0.054147	0.059	1.094255
32	118.2	4.772	-0.01428	0.054106	0.010	0.18669
33	119.4	4.782	0.039822	0.054226	0.101	1.86404
34	132.1	4.884	0.094048	0.054519	0.056	1.026031
35	139.7	4.939	0.148567			

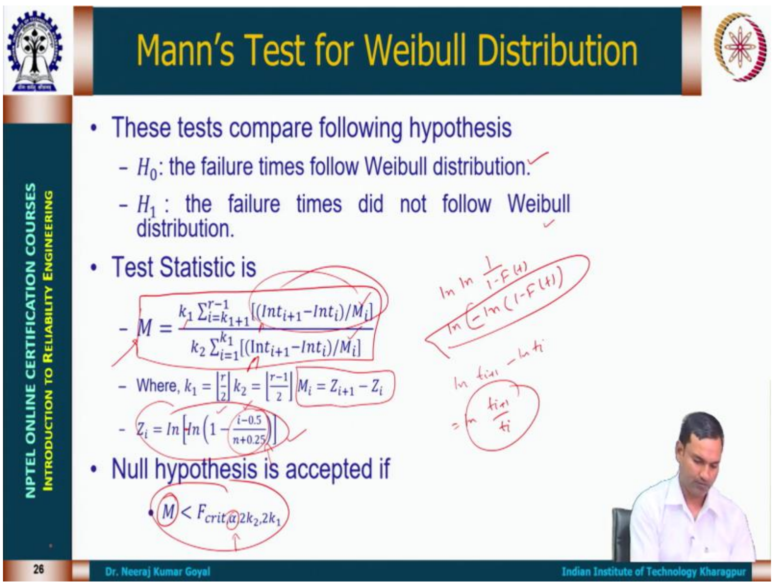
n	50
r	35
k1	17
k2	17

M	1.621544
Fcrit	1.772066

$M < F_{crit,\alpha,2k_2,2k_1}$

Null hypothesis is accepted





Test Statistic is

$$M = \frac{k_1 \sum_{i=k_1+1}^{r-1} [(\text{Int } t_{i+1} - \text{Int}_i)/M_i]}{k_2 \sum_{i=1}^{k_1} [(\text{In } t_{i+1} - \text{Int } t_i)/M_i]}$$

Where, $k_1 = \left\lfloor \frac{r}{2} \right\rfloor k_2 = \left\lfloor \frac{r-1}{2} \right\rfloor M_i = Z_{i+1} - Z_i$

$$Z_i = \text{In}\left[\text{In}\left(1 - \frac{i-0.5}{n+0.25}\right)\right]$$

So, let us say if we have the exam, let us take an example, that we have the 35 failure points here, out of 50 we had put 50 devices on test out of his 35 failed and for 35 this is our data and so, 35 divided by 2 will be 17.5 and floor value of this will be 17. So, that will be my k 1 and k 2 is equal to r minus 1 that means 34 divided by 2, 17 floor value of 17 is 17only.

So, my k 1 and k 2 both turns out to be 17. From this I can calculate the value of M my purpose is to calculate M. So, as I see here to calculate M, what are the things I need to know if you see I have to take a summation of i equal to 1 to k 1 plus 1 to r minus 1 and divided by i equal to 1 to k 1 and here what are the values I am using ln of ti minus 1, minus ln of ti ln of ti plus 1 minus ln of ti. So, I have to calculate ln of ti plus 1 I have to calculate ln of ti and this has to be divided by Mi, and Mi is what, Mi for Zi plus 1 minus Zi.

So, if I measure Z if I calculate ln of t, I will be able to use this so what I have done we covered basic data is In ti these 2 i already have so what I did I calculate the ln of t i because I know I need to have the ln of ti and ti plus 1 is the next value for i plus 1. But and next value is Zi, Zi is nothing but the probability. So, this test we are doing for Weibull distribution, so for Weibull distribution we can get this Zi, Zi is ln of minus ln 1 minus F t. Sorry, and F t is i minus 0.5 divided by n plus.

So, we will apply the same formula i minus 0.5 divided by n plus point so, i is known to us because Z is function of only i. So, if I know the i, I will know the Z same formula we have applied I have done this in Excel sheet which we will share and you will get this values. So, Zi values are known to us. Now, Mi values also we can get. What is Mi? Zi plus 1 minus Zi.

So, that means this minus this will give me this. So, these values I will get by subtracting from. So, I will get one value less here because I have to subtract this value from this so, one value will be less.

Similarly, ln of ti plus 1 minus ln of ti. So, this ln of ti plus 1 minus ln of ti we can write it as ln of ti plus 1 divided by ti. So, that I that I can subtract or I can do like this also. So, ln of ti plus 1 this is ti plus 1 divided by ti. So, once that means or I can take the subtraction so, this is this minus this if I take I will get this 1.988 minus 1.262 will give 1.726.

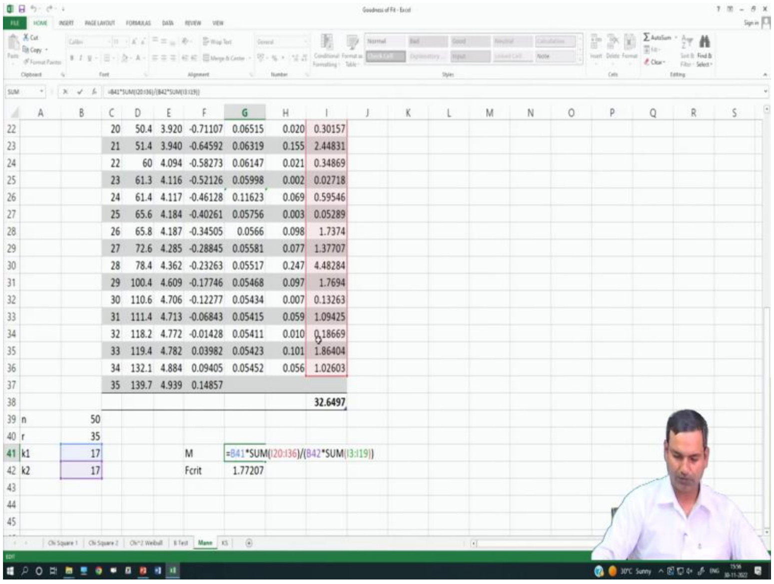
So, this is subtraction from this I can get subtraction also or I can simply take division also ti division and take the ln whatever the will it will give you the same value. Similarly, now, I have got these values. So, these values divided by Mi here also and here also. So, I will divide this by Mi, Mi is here this value is here, so, this becomes my value.

So, here if you see now, I have calculated this whole value. So, this whole value I have calculated here, so, if I take summation of this, now, I have to take the summation in 2 parts, one is from k 1 plus 1 to r minus 1 and another is i equal to 1 to k1. So, 1 to k1, k 1 is 17. That means one set is here and another set is from here to here. Since I have only 34 data points r minus n, r is 35 but I am taking only up to 34.

So, two 17, 17 data sets I have, I will take the summation here, I will take the summation here these 2 summations I will calculate and these 2 summations then we are comparing. So, from k plus 1 to r minus 1 that means this summation and this is i equal to 1 to k1 that is the summation. So, this summation let us say this is S2 and this is S1. So, I will take S2 upon S1 and S2 is multiplied by k 1 and as soon as multiply by k 2, so, this is k 1 k 2 once I do this I will get the value of n.

So, this simple formula I have applied in Excel and I have calculated this value of M. And how do I get the critical value of F. So, critical value of f is by taking the F distribution and taking this inverse of F distribution for right side values that is alpha 2k 2 and 2k 1. Degrees of freedom for F distributions are 2k 2 and 2k 1, 34, 34. So, once I use this, I will be able to this I have done here also same thing.

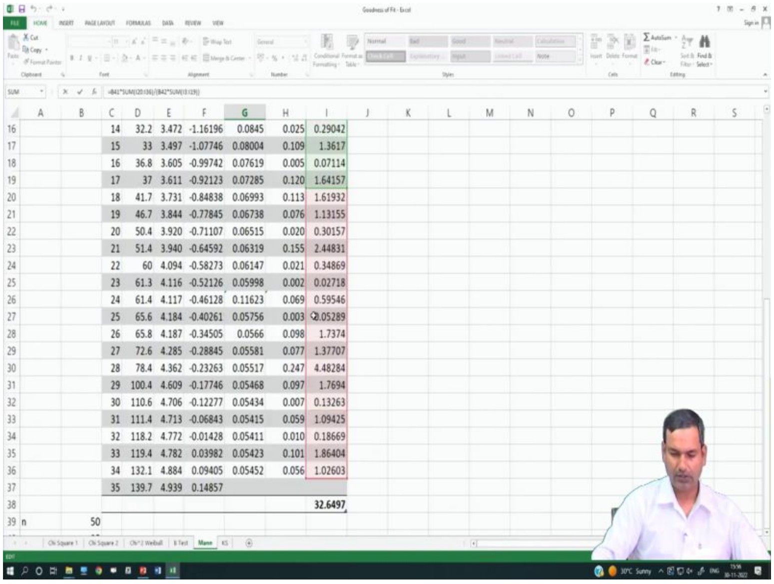
(Refer Slide Time: 29:03)



Goodness of Fit - Excel

=B41*SUM(I20:I36)/(B42*SUM(I3:I19))

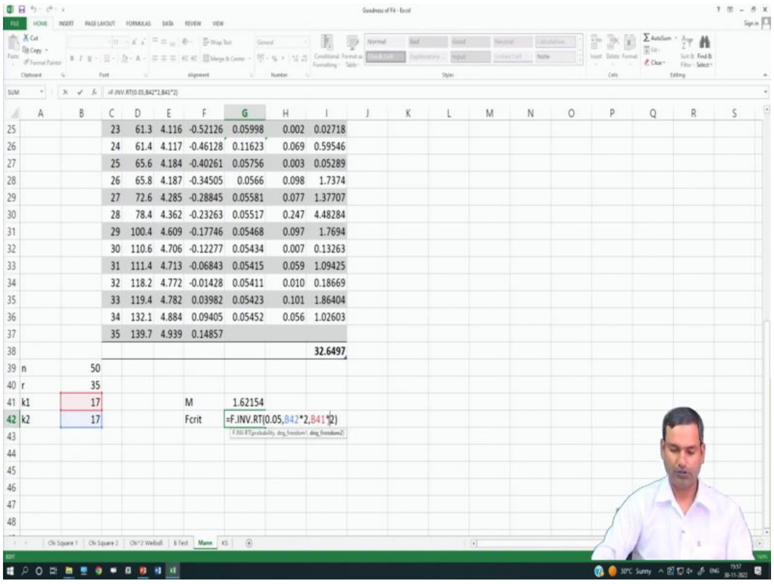
	A	B	C	D	E	F	G	H	I
22			20	50.4	3.920	-0.71107	0.06515	0.020	0.30157
23			21	51.4	3.940	-0.64592	0.06319	0.155	2.44831
24			22	60	4.094	-0.58273	0.06147	0.021	0.34869
25			23	61.3	4.116	-0.52126	0.05998	0.002	0.02718
26			24	61.4	4.117	-0.46128	0.11623	0.069	0.59546
27			25	65.6	4.184	-0.40261	0.05756	0.003	0.05289
28			26	65.8	4.187	-0.34505	0.0566	0.098	1.7374
29			27	72.6	4.285	-0.28845	0.05581	0.077	1.37707
30			28	78.4	4.362	-0.23263	0.05517	0.247	4.48284
31			29	100.4	4.609	-0.17746	0.05468	0.097	1.7694
32			30	110.6	4.706	-0.12277	0.05434	0.007	0.13263
33			31	111.4	4.713	-0.06843	0.05415	0.059	1.09425
34			32	118.2	4.772	-0.01428	0.05411	0.010	0.18669
35			33	119.4	4.782	0.03982	0.05423	0.101	1.86404
36			34	132.1	4.884	0.09405	0.05452	0.056	1.02603
37			35	139.7	4.939	0.14857			
38									32.6497
39	n	50							
40	r	35							
41	k1	17				M	=B41*SUM(I20:I36)/(B42*SUM(I3:I19))		
42	k2	17				Fcrit	1.77207		



Goodness of Fit - Excel

=B41*SUM(I20:I36)/(B42*SUM(I3:I19))

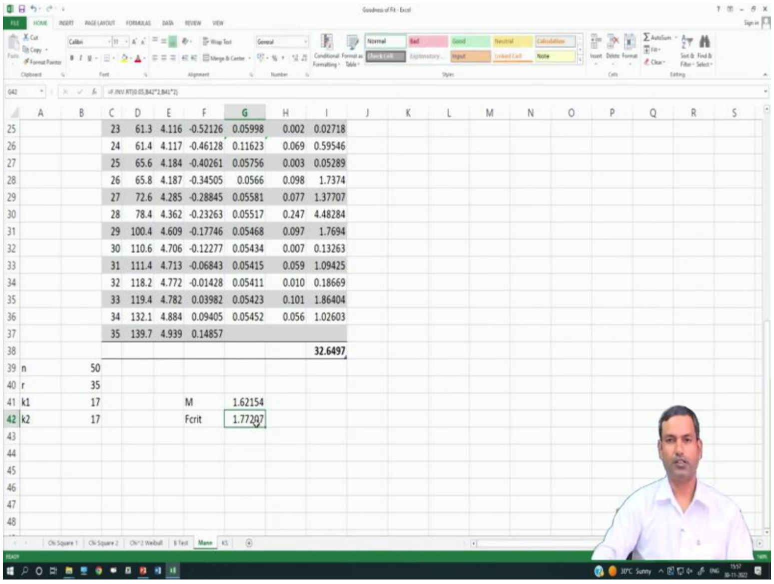
	A	B	C	D	E	F	G	H	I
16			14	32.2	3.472	-1.16196	0.0845	0.025	0.29042
17			15	33	3.497	-1.07746	0.08004	0.109	1.3617
18			16	36.8	3.605	-0.99742	0.07619	0.005	0.07114
19			17	37	3.611	-0.92123	0.07285	0.120	1.64157
20			18	41.7	3.731	-0.84838	0.06993	0.113	1.61932
21			19	46.7	3.844	-0.77845	0.06738	0.076	1.13155
22			20	50.4	3.920	-0.71107	0.06515	0.020	0.30157
23			21	51.4	3.940	-0.64592	0.06319	0.155	2.44831
24			22	60	4.094	-0.58273	0.06147	0.021	0.34869
25			23	61.3	4.116	-0.52126	0.05998	0.002	0.02718
26			24	61.4	4.117	-0.46128	0.11623	0.069	0.59546
27			25	65.6	4.184	-0.40261	0.05756	0.003	0.05289
28			26	65.8	4.187	-0.34505	0.0566	0.098	1.7374
29			27	72.6	4.285	-0.28845	0.05581	0.077	1.37707
30			28	78.4	4.362	-0.23263	0.05517	0.247	4.48284
31			29	100.4	4.609	-0.17746	0.05468	0.097	1.7694
32			30	110.6	4.706	-0.12277	0.05434	0.007	0.13263
33			31	111.4	4.713	-0.06843	0.05415	0.059	1.09425
34			32	118.2	4.772	-0.01428	0.05411	0.010	0.18669
35			33	119.4	4.782	0.03982	0.05423	0.101	1.86404
36			34	132.1	4.884	0.09405	0.05452	0.056	1.02603
37			35	139.7	4.939	0.14857			
38									32.6497
39	n	50							



Goodness of Fit - Excel

=F.INV.RT(0.05,B42*2,B41*2)

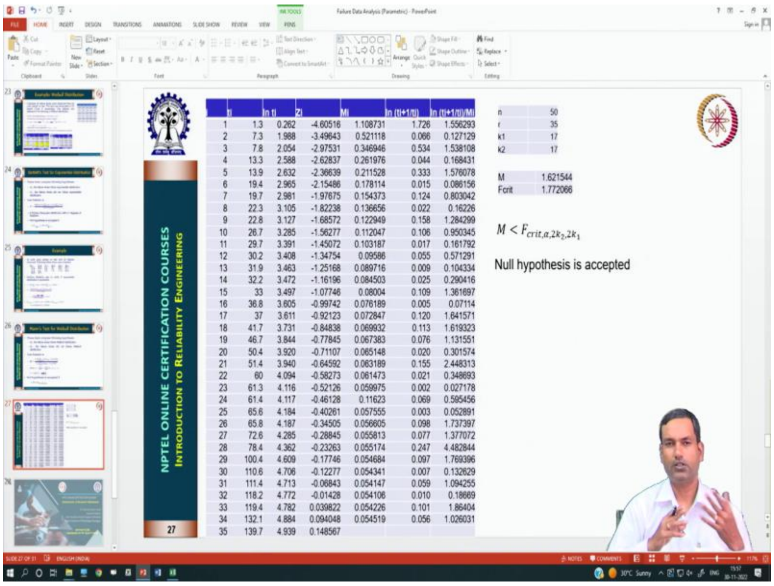
	A	B	C	D	E	F	G	H	I
25			23	61.3	4.116	-0.52126	0.05998	0.002	0.02718
26			24	61.4	4.117	-0.46128	0.11623	0.069	0.59546
27			25	65.6	4.184	-0.40261	0.05756	0.003	0.05289
28			26	65.8	4.187	-0.34505	0.0566	0.098	1.7374
29			27	72.6	4.285	-0.28845	0.05581	0.077	1.37707
30			28	78.4	4.362	-0.23263	0.05517	0.247	4.48284
31			29	100.4	4.609	-0.17746	0.05468	0.097	1.7694
32			30	110.6	4.706	-0.12277	0.05434	0.007	0.13263
33			31	111.4	4.713	-0.06843	0.05415	0.059	1.09425
34			32	118.2	4.772	-0.01428	0.05411	0.010	0.18669
35			33	119.4	4.782	0.03982	0.05423	0.101	1.86404
36			34	132.1	4.884	0.09405	0.05452	0.056	1.02603
37			35	139.7	4.939	0.14857			
38									32.6497
39	n	50							
40	r	35							
41	k1	17				M	1.62154		
42	k2	17				Fcrit	=F.INV.RT(0.05,B42*2,B41*2)		



So, this same table which I have shown here there, it is the same and if you see here, how do I calculate M, M is k1 into summation of the second part red part and divided by k2 that is B 42 and the green part that is I3 and this when I take this gives me the M and F critical are how I have got that is the F inverse rightside values for this I have taken for 0.05 that means 5 percent, 95 percent confidence level.

If I want I can take the 10 percent also or anything else and my degrees of freedom is k2 into 2 and k 1 into 2 and I use my critical value is 1.72. So my M value is lesser than critical value. So, I can say that but it is quite close. So, but still I can say that my Weibull distribution is justified here for this use.

(Refer Slide Time: 30:12)



So, here as we discussed today that we can have specific checks we can find out whether a particular distribution is followed or not based on general tests like chi-square test or we can also use a specific test designed for the purpose. So, thank you. We will continue our discussion in next class.