Introduction to Reliability Engineering

Professor Neeraj Kumar Goyal
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

Lecture 31

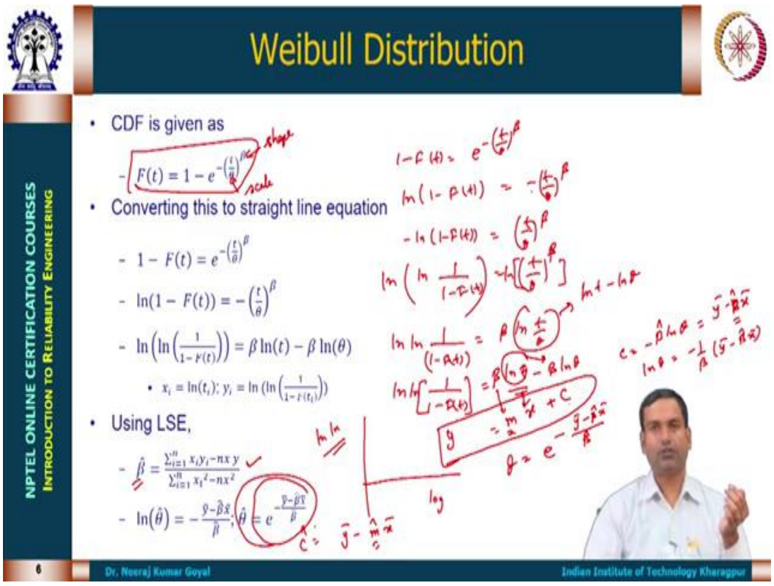
Failure Data Analysis (Parametric) (Contd.)

(Refer Slide Time: 00:26)

Hello everyone, we will continue our discussion from where we left in previous lecture, we there we discussed exponential distribution, how we can fit the exponential distribution to know its parameter, because if we know some data is fitting exponential distribution, we just need to know the parameter of exponential distribution, which is the lambda. Once we know the lambda value, then we know so, many components may follow the exponential distribution, but they may have a different lambda value.

And based on that, the reliability would be different same way the there may be different distributions also, they may also follow a log normal distribution or Weibull distribution. So, let us say if we have the data and we want to fit to the Weibull distribution, then how we will fit the data the same approach which we use for the exponential distribution same we will use in Weibull also, but there will be little change in formula because Weibull distribution has a different formula, so different pattern is to be captured.

(Refer Slide Time: 01:31)



CDF is given as: $F(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^{\beta}}$

Converting this to straight line equation

$$1 - F(t) = e^{-\left(\frac{t}{\theta}\right)^{\beta}}$$

$$\ln(1 - F(t)) = -\left(\frac{t}{\theta}\right)^{\beta}$$

$$\ln\left(\ln\left(\frac{1}{1 - F(t)}\right)\right) = \beta \ln(t) - \beta \ln(\theta)$$

$$x_i = \ln(t_i); y_i = \ln\left(\ln\left(\frac{1}{1 - F(t_i)}\right)\right)$$

Using LSE,

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2}$$

$$\ln(\hat{\theta}) = -\frac{\bar{y} - \hat{\beta}\bar{x}}{\hat{\beta}}; \hat{\theta} = e^{-\frac{\bar{y} - \hat{\beta}\bar{x}}{\hat{\beta}}}$$

So, let us see how do we go with the Weibull distribution. So, for Weibull distribution our CDF formula, as we know is given as ft is equal to 1 minus e to the minus t upon theta raised to the power beta where theta is the scale parameter and beta is the shape parameter. This we have discussed in earlier classes. So, a scale parameter defines the scale if larger that means the spread is larger on larger values and beta changes the shape of the distribution.

Now, we want to fit this to the straight line equation. To do that, what we will do first we will take 1 will do 1 minus ft 1 minus ft will become e to the minus t upon theta raise to the beta now, what we will do we will take ln of both sides we will take ln then what we will get is minus t upon theta is to the beta as you see here, if we take single n ln then it is not sufficient enough because the t is still not coming out directly as the function still t raise to power beta is coming.

So, we have to do 1 more ln, but before that, because this is minus sign we have to make it positive. So, we will take minus ln of 1 minus ft here and that will become T upon theta raise to the beta this minus ln I can also write it ln of 1 upon 1 minus ft that is equal to t upon theta raise to the power beta now, here I can take another log again and log of this what will happen this will become ln ln of 1 upon 1 minus ft this will be equal to log of this so, log of this power means power will come out so, this one beta ln of t upon theta and ln of t upon theta is nothing what ln of t minus ln of theta.

So, this will become beta ln of t minus beta ln of theta and this is my ln double log of 1 upon 1 minus ft as we see here that we actually when we want to have the graph paper for this equivalent for exponential you need to have a semi log where 1 accesses log and other accesses is the linear. But here as you see that time is also converted into log value.

So, x axis also has to be logarithmic or natural logarithmic and Y axis also you need to have the double log paper. So, if you want to plot it on the graph paper and see the linearity in the values, you have to use the log log paper log log paper on y axis and log on x axis once you use that kind of paper, then only this you will have the linear equation.

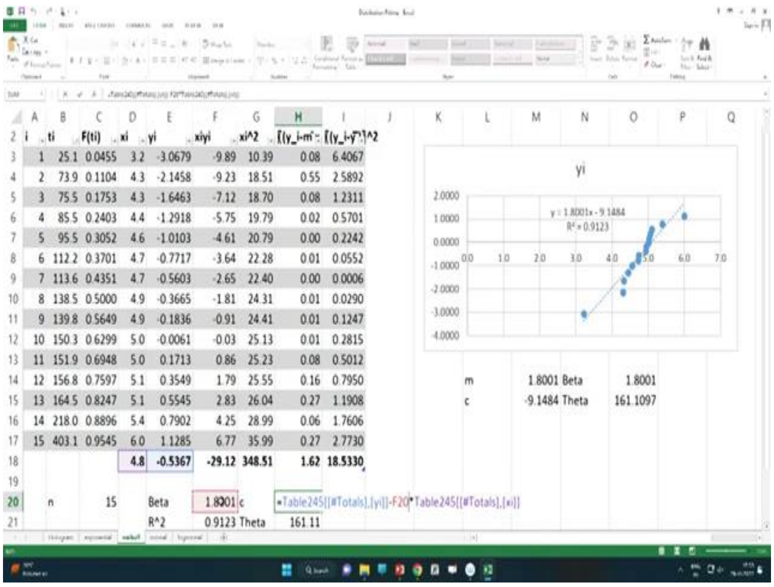
So, here this is y and this is beta is m x is ln of T and c is minus b ln of theta. So, same equation we can use what we have used earlier same regression formula we can use and we will get this. So, if you want to know the beta value which is m. So, whatever the formula of m which you develop that is summation of i equal to 1 to n xi yi minus n xi x bar y bar divided by summation of i equal to 1 n 1 to n xi squared minus and x bar squared.

Same formula we will use here for this value of c we know what is the value of C, c cap is equal to as we see here that is y bar minus estimated value of m minus x bar this is what we have seen earlier that value of c is equal to here y bar minus mx bar same value we can use here y bar minus mx bar here y bar is the average value of y x bar is average value of x which we will get from the data and m bar we will get from here, which is the beta cap.

So, if I and what is c cap here c is equal to minus beta ln of theta. So, I want to know the value of theta beta value I have already got and this is equal to y bar minus beta x bar, because m is equal to beta. So, now, I want to get so, this will ln theta will become minus 1 upon beta y bar minus beta cap x bar.

And from here if I want to know that theta theta will be exponential of this. So, theta will be equal to e to the power minus y bar minus beta cap x bar divided by beta cap. So, same value we will be able to get the theta. So, here what we have to do we have to we have converted the equation to the line equation. So, first we calculate the xi yi values, then we calculate the value of m and c this I have shown here for the data I will show it in the Excel sheet so, that you are able to follow it properly.

(Refer Slide Time: 06:54)



So, I have developed the same thing in the Excel sheet let us see this is Weibull. So, let us say we have the 15 time to fill a data, which I have already sorted and put it here. So, i equal to 1 to ti I have this 50 data and Fti as we discussed earlier Fti is nothing but i minus 0.3 divided by n plus 0.4. So, same formula that we have used earlier same formula is used again i is if you see here a3 is i minus 0.3 divided by n value I have put it here, that is 15 minus 15 plus 0.4.

So, this value will give me the Fti values, I will get the Fti values. Now, I know, for ti and Fti, I have to convert it into the xi and yi. And what is xi here as we see here, xi is nothing but the ln of ti, so what I will do in this sheet also, for xi calculation, I will take the log of ln of ti, I will do that, and I will get the xi value.

So, time I have converted into xi. Similarly I will get the yi, what is yi, yi is double log of yi is the double log of 1 upon 1 minus f ti, or we can see log of minus ln 1 minus f ti whatever the way we can calculate. So same thing I have already calculated, I will put it again to show you ln of I can take minus ln so that I do not have to write 1 upon 1 minus f ti or I can take let us say ln of ln again ln of ln 1 divided by 1 minus Fti, Fti is this value. And this becomes my values for the yi.

Now, I know the xi and yi. So, if I use the approach like earlier, I have already plotted here, but as for doing it again, for clarity, I will do it again. What I will do, I will plot it again, I will do go for insert first let me select the data once again, xi versus yi, I want to insert a plot, what is the plot type plot type is the xy scatter. So, I will use the xy scatter here. If you see I will get the xy scatter diagram here.

Now next step as you remember what we have to do we have to fit the trendline to this and which type of trendline, we have to fit we have to fit the linear graph here, linear sink's straight line fit. For exponential distribution we had to set the intercept 0 but for Weibull distribution we do not have to set the intercept to the 0 because Weibull distribution has a intercept.

So, intercept has to be calculated, we will display the equation and we will display the R squared value once we do that we are able to get the fit to the straight line if you see that if we look at here my m value is coming out to be 1.8001 and how much is my c value c value is equal to minus of 9.1484 from earlier discussion we know that what is my m, m value same as beta. So, this value what I am getting as m this value is also the beta value. Now, I want to calculate the theta value.

So, beta value is equal to same as m and I want to calculate theta value theta value as we have discussed earlier theta value is equal to exponential of minus of minus of y bar minus beta x bar divided by beta, where is y bar this yi I have averaged here this is my y bar, this is the average value of y, y bar minus beta this is beta into x bar x bar also I have calculated here average value of x x bar divided by beta divided by beta once I do this I get the value of theta theta is the characteristic life.

If I want I can also calculate the theta from here. How can I calculate theta from here, as you remember that at whenever I put t equal to theta in this equation, if I put t equal to theta this will again become 1 minus e to the power minus 1 which is similar to exponential distribution that means 62.2 percent or on log scale it was 1.

So, log of log will become 0 so, that means, wherever it is crossing 0 that will be your value. So, somewhere here 5 but this is 5 is on here is the log values. So, exponential 5 you have to take. So, better to take calculated directly because otherwise I have to plot rather than t I have, ln of t I have to first calculate the t value here corresponding t value so, that intercept at 0 somewhere here around 5.1 something so, exponential of 5, 5.1 If I take exponential of 5.1 somewhere around you will get the value 161 164 because I am not able to accurately read it so, that is why this difference is coming.

So, here we are able to get the beta and theta from the graph. Now, let us see the approach which we discussed based on the statistics. So, for statistics as we discussed earlier what we have to do we can use this formula beta is equal to summation of xi xi yi minus n into x bar y bar.

So, like we have done for exponential I have calculated the xi yi here that is nothing but multiplication of xi and yi this is equal to xi into yi if I do that, I will get this all this value and once I get this xi yi then another value in denominator what I need is summation of xi square x bar and y bar we will get from the bottom of the table so, need not to calculate it again. What is x bar here? This is my x bar average value of xi, and what is y bar here that is average value of y bar yi.

So, and for denominator I need to take the summation of xi square so what I will do I will first calculate the xi square here. So, xi square is this is equal to xi power 2, and we are able to get the xi square here. And summation of xi squared is we will take the total sum of this. So, now we have all the value we have the summation of xi yi, this is sum.

So, how much is beta beta is equal to I can calculate it again for your reference that is equal to as we see, we have the summation of xi yi. This is my summation of xi yi minus n into x bar y bar this is my x bar this is my y bar divided by summation of xi squared. So xi squared summation is this minus n into x bar squared.

So, x bar is this x bar square we get the same here because decimal 0.91 0.9123. So, we are able to get 1.8, so we are able to get the same value here actually decimal points are shown are less I will increase the decimal point for your reference so that you can see 1.8001, so we got the beta value n values is beta so I am able to calculate that directly.

And what is my R squared value, R squared value again, I will use this formula which we have seen earlier. R squared is 1 minus summation of y minus mx i minus Ci whole squared divided by yi minus y bar whole square.

So, same thing I will do it here. This is already done 1 minus this is my yi, mx I will do this again let me do this again. This is equal to 1 minus y minus so summation of y, so for this we have already calculated if you see here, I have already calculated y minus mx i minus c. So, this is equal to let us see this again y minus c minus mxi.

So, same thing I have done that is equal to y, this is my y minus c mxi. So, xi is this multiply by m value of m I have already calculated here, so, I will put, since I do not want this to change, column wise, so I will put the dollar sign here m minus c value of c also have already calculated that is, I do not want to change again so, I will put this dollar sign again.

Once you, I do that, this becomes my yi value. And some of this value yi value I can get it here but this value, I have to take the square of this so I will take this square. So, I will do this square here itself. As you have seen that, if you do not do this square error sum is 0, which is natural positive, because we want to fit the trendline which is minimizing the error. So, approximately, it will try to make the positive and negative errors equal to 0. So, here, we are able to get this.

And we want to calculate yi minus y bar square. So, this is equal to yi minus y bar y bar means average value of y this is my average value y and whole square of this. So, c value we have already calculated c value is, that is y minus if you know the formula for c value y minus mx.

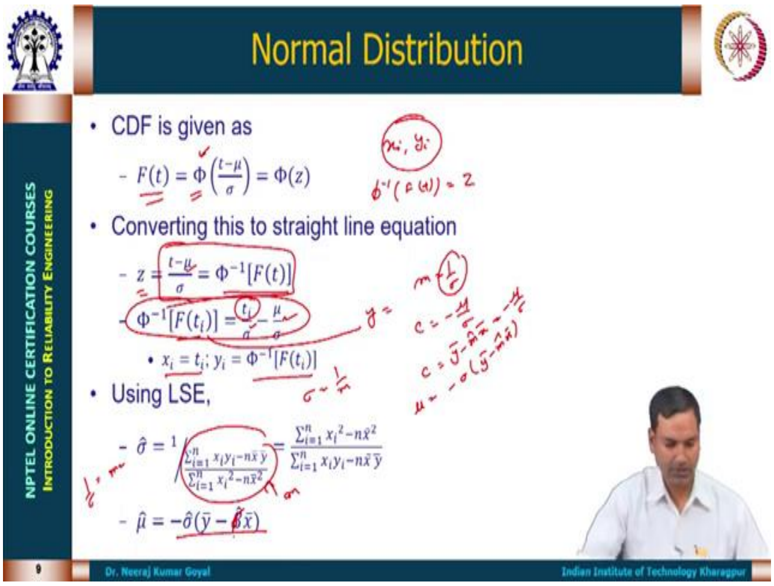
So, y value average y values or I can do this again. This is equal to y average y minus m is this into x average x average is this into x average we get this. So, similarly, we are able to get, I think, did I make some error here, y y minus cmxi y minus mxi minus C, c is equal to y bar minus x bar into beta.

So, once we are able to get this then theta is let me just check where did I make the mistake, value of c is minus 9.95 which is yi bar minus f 20 f 20 beta into same thing, this is same, and then we get this this values yi minus xi into beta squared, maybe I have made some error there. So, that is where the value got changed.

So, once we have this we got a beta value we got the R squared value we got the c value and from the c value I can calculate that theta again using the same way same way which we have calculated if you see this is the same value minus 9.9418 minus 9.1484 and theta value again I can calculate the same value that is equal to exponential minus of minus of see this minus of y minus beta x divided by beta y minus beta into x divided by beta 161.11.

So, we are able to calculate the values using this calculations also or we can use the Excel function for the trendline fitting also, which will also give you the same result. So, this we are able to do for the Weibull distribution, we are able to do the exponential distribution let us see if you want to do this for other distributions like if same thing are plotted here.

(Refer Slide Time: 23:06)



Now, let us see if we go for the normal distribution. So, for normal distribution, what is my ft ft is given in terms of now only thing we have to change here is this ft value the only equation needs to be changed, but final formula and everything is going to be same. The only thing is how do I calculate yi and xi that will only change from value of ti and Fti I have to calculate the xi and yi.

Now, for normal distribution we know cumulative distribution ft is given as a standard normal value phi cumulative value phi for z z is t minus mu divided by sigma. So, here what I can do I can take the phi inverse phi inverse of ft that will give me a z. So, z is equal to phi inverse of ft and what is z t minus mu divided by sigma.

CDF is given as

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma}\right) = \Phi(z)$$

Converting this to straight line equation

$$Z = \frac{t-\mu}{\sigma} = \Phi^{-1}[F(t)]$$

$$\Phi^{-1}[F(t_i)] = \frac{t_i}{\sigma} - \frac{\mu}{\sigma}$$

$$x_i = t_i; y_i = \Phi^{-1}[F(t_i)]$$

Using LSE,

$$\hat{\sigma} = 1/\frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2} = \frac{\sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2}{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}$$

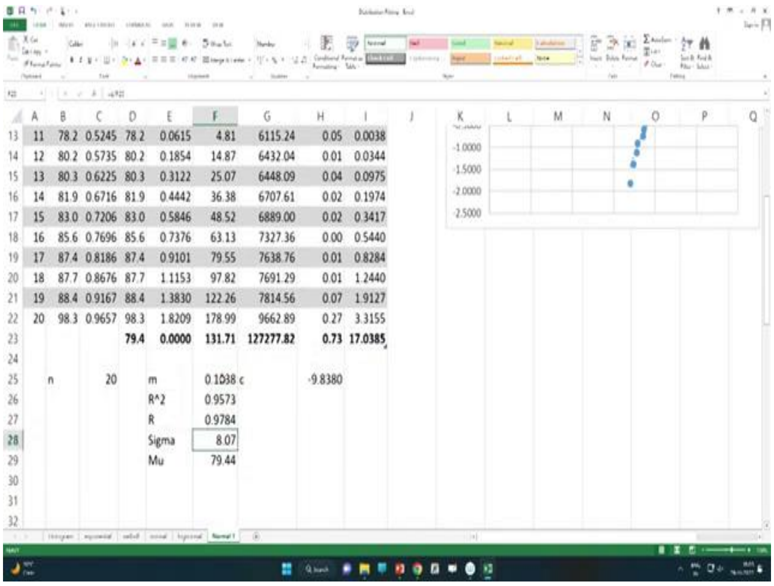
$$\hat{\mu} = -\hat{\sigma}(\bar{y} - \hat{\beta}\bar{x})$$

So, from here this I can write if I write the individual ti value this will be phi inverse of Fti that will be equal to ti divided by sigma minus mu divided by sigma. So, this becomes my straight line where y is equal to phi inverse Fti and xi is remain same ti I do not have to take double ln here.

So, single xi is ti and m is equal to 1 upon sigma and c is equal to minus mu divided by sigma. Once I have done this, then I can use the same formula same sheet by making little changes there and this formula is same whatever my m is here. So, m is equal to 1 upon sigma. So, this value was for the m.

So, my sigma value will be 1 upon m here because m is 1 upon sigma, same way whatever I get the value of c, so c values minus mu upon sigma. So, and c is as we discussed c is y bar minus m cap x bar. So, and this is equal to minus mu divided by sigma. So, my mu value will be again equal to minus sigma into y bar minus m cap x bar minus sigma into y bar minus beta x bar it will not be beta it is sigma. So, since there is a type typographical error here, so, this will be sigma x bar. So, we can do this same.

(Refer Slide Time: 25:43)



So, let us see if we have the data, how do we do it. So, here I have done this, one way of doing this is like I can copy this whole sheet again, what I can do, let me show you in a way that how I generally do this let us say create a copy and move to the end. If you see I have got let us say, normal distribution here I want to fit it to the normal distribution.

Now, I have the data already here. I have 20 filler data here, but this table was for 15. So, what I will do, I will insert 5 more here. I will insert 5 more rows here, once I insert 5 more rows now, what will I do is I will take the this data already let us say I have this is the failure data which I have observed for normal distribution, that means I have the 20 filler data points here and for 20 data points I want to fit to fit to the normal distribution.

So, what I will do I will this this is my basic data Inti whatever is IntIi based on that my Fti etcetera has to be calculated. Now, Fti value as we have to calculate see everything else xi yi xi square y this all is going to be same formula going to be same, what will be the change, change will be in xi yi and Fti value will also be same Fti value same formula will be applicable, the only change will be this n will become 20.

So, for 20 because now I have 20 data points, so n has become 20 same formula will be applicable here, and I will get the Fti. Now, xi value how I have to do change, xi value here is t, so but I will do xi is equal to ti. Now, what is yi value here yi values the phi inverse Fti. That means what I have to do I have to take the norm inverse value.

So, Excel has the function for normal inverse. So, that is norm dot inverse. So, I will use norm inverse value, the probability for which I want to get this value is Fti. And this is the standard normal distribution I want to use.

So, for a standard normal distribution the mean is 0 and standard deviation is 1. With this I get the new yi once I get the new yi everything is I do not have to calculate xi yi and other things again, because that is by default same for same formula based on the xi yi value is will be applicable and same way I will get it, but what I will get here is the c value, this is the my 1 upon sigma or this I will call it as this is my m value and my c value also is going to be same and m value is going to same from here, I can now calculate the beta and theta.

So, rather than beta and theta, but I will calculate is I will calculate the sigma. So, sigma as we know sigma is equal to 1 divided by m. So, this will become my sigma and mu mean value mu is equal to sigma is standard deviation also and value of mu as from this formula mu is minus sigma y bar minus beta x bar.

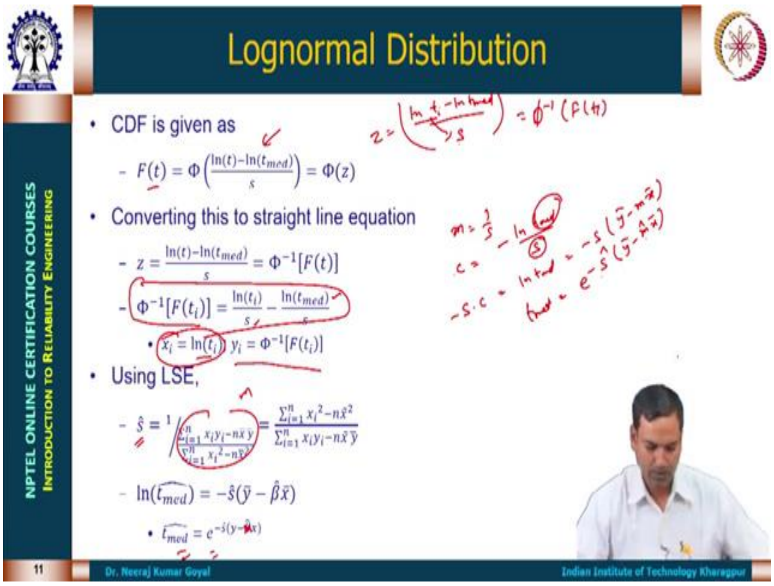
So, that is equal to minus sigma into y bar my summation of y is this y bar yi bar is this y bar minus mx bar m into x bar. This is my x bar so, my mu becomes 79.44 y minus this will not be beta this will be m mx bar.

So, we will be able to use the same function here and this mu comes out to be 79.44 if you see it here, this m 1.123 x 1.1238 and c is point minus 9.838 minus 9.838, I do not need theta here, I will not use it. And R squared value again is 0.9573 I have calculated, calculated values also 0.9573.

So, as you see here, by using that because the approach is same, the formulas basic formula for calculation of m and c is same, but the meaning of m and meaning of c may be different from distribution to distribution.

And because of that, when we are using different distribution, the distribution parameter I have to calculate differently like sigma is 1 upon m here, but for Weibull beta was equal to m for exponential the lambda was equal to m, but here this is 1 upon m and similarly, mu also I have a formula that minus sigma y bar minus m x bar. So, same formula I have used here. So, same way we can solve like same thing if we look at here if we have the this we have done for normal distribution.

(Refer Slide Time: 31:30)



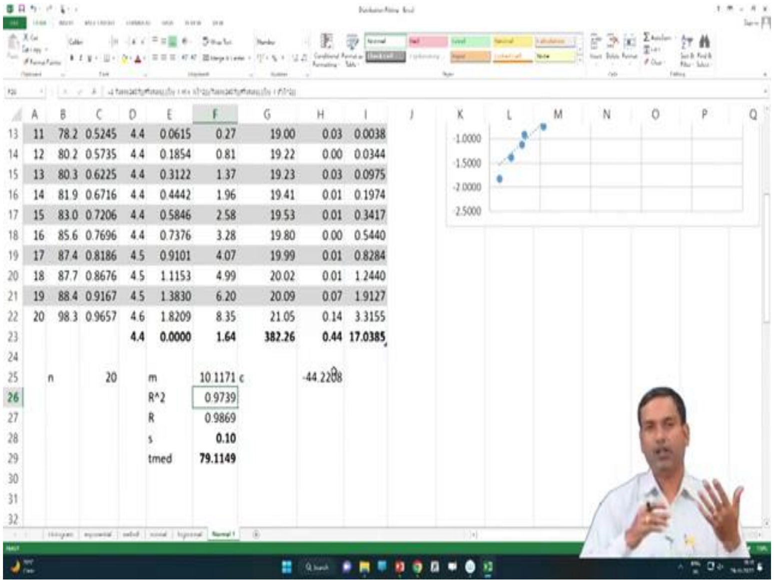
If log normal distribution is very much similar to the normal distribution, the only difference which we have in log normal and normal is like ft formula, because z will become ln of t minus ln of t median divided by s as we know from the log normal distribution z is given like this and this is equal to phi inverse of Fti.

So, same way I can use this value this function so, my phi inverse Fti will be equal to ln of ti divided by s minus ln of t median divided by s if I take the 2 parameters t median and s this will be my formula, ln of ti divided by s and ln of t median. Now, what the changes happened for normal distribution it was i, but for log normal distribution xi will be ln of ti and yi is same phi inverse of Fti another change would be that earlier we were calculating my m value which would be here 1 upon s and c value what we what I have here is minus ln of t median divided by s.

So, once I calculate the m so, 1 upon m will give me the s and from here once I get the S value I can calculate the t median, t median would be going to the e to the power minus this if I say this will be s into c minus would be equal to ln of t median and how much is c c is equal to minus s into y bar minus m x bar.

So, this again this beta is written this is m, m x bar. So, t median is equal to e to the power minus s into y bar minus mx bar where m is estimated value and s is also estimated value. So, same way I can use, m I can also put it as 1 upon s, similarly this s I can put it 1 upon m, both is going to be same. So, the same thing like I have done it already for log normal distribution.

(Refer Slide Time: 33:50)



So, for log normal distribution if I if you see here, same sheet are copied what changed I have to do it here like the same sheet if I am going to use here let us say I fit the same data to the log normal distribution, what changes I have to do is first 3 columns are going to be same there will be no change because this is determined by the data i ti and Fti what changes there xi becomes ln of ti.

So, I will make this ln of ti and what is my y y remains same as as we discussed earlier y for normal distribution and for log normal distribution that is same that is phi inverse of of Fti phi inverse of 1 minus Fti so, same thing is there, I think there is one mistake here phi inverse of 1 minus Fti we have to take, no no phi inverse of Fti only.

So, phi inverse of Fti we take, we got this and same rest of the things are saying y xi xi square this is all remains same as you see here, this is my log normal fit now, the earlier it was the normal fit, but now this is log normal fit.

So, for normal log normal fit when I am doing then m is equal to 1 upon s, so, this will become s, s will be equal to 1 upon n c is this. So, I want to calculate the t median here. So, t median would be equal to as we see we will use this formula which you have seen t median is e to the power minus s into that is exponential minus s value of s multiply by y bar this is my y bar minus m into t bar, m into x bar m into x bar.

So, this will give me the t median that is 80. So, we are able to get the parameters s and t median for the log normal distribution for the same data and if you see the fit value this is

0.9739, while earlier this was 0.95573. So, if you see here my fit is better compared to normal distribution for log normal distribution the fit value was 0.9573 r square and r square value is 97739 here, which is greater than earlier value 9573.

So, that means, better the r square better is the fit. So, we can see that this data is fitting better to the log normal distribution. And we can assume this data is fitting to the log normal distribution rather than normal distribution.

And once we assume that, we can now use the parameters of log normal distribution, we will know the reliability we will know the failure rate we will know zt ft everything as we discussed in when we discuss the when we discussed this distribution in initial classes maybe in second or third week.

So, when we discussed that, so, that once we get the parameter from here then whatever discussion we did in earlier weeks. So, the values comes from here from the data analysis and that is used over there.

And once you are using over there, you will be able to get all the important decisions making all the important reliability characteristics or unreliable characteristics or even here we are discussing the time to failure rather than time to failure if we consider the data time to repair, then nothing changes the distribution would be for the time to repair except that everything will be same. The only the variable name will change and we can calculate the maintainability there, we can calculate MTTR rather than MTTF it will become MTTR, and we will be able to use these values as we require in various analysis.

So, with this like we have discussed about distribution fitting. We have discussed four distributions here exponential, Weibull, normal log normal, which are commonly used distributions most commonly used distribution in the analysis of reliability, that is time to failure data, as well as the time to repair data, which is the major data for the reliability and maintainability analysis or the availability analysis. So, we will stop here today and we will continue our discussion more about this distribution fitting and more about the maintainability availability etcetera in coming lectures. Thank you.