

Introduction to Reliability Engineering
Professor Neeraj Kumar Goyal
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur
Lecture: 30
Failure Data Analysis (Parametric)

(Refer Slide Time: 00:26)




Hello everyone. So, in previous sessions we discussed about failure data analysis that is mostly time to failure analysis or we have the group failure data analysis, which may be completed or which may be group data or sensor data. So, and we discussed how we can get the user rank methods and we are able to get the $F(t)$ once we get the $F(t)$, then we can fit it to we can find out the $R(t)$ or $Z(t)$ or $f(t)$ like that.


So, today onwards we will be discussing more about how we can use the same data and we are able to fit the distribution. Once we fit the distribution we are able to know the distribution pattern, once you know the distribution pattern, we will be able to evaluate important quantities related with the reliability as well as we will be knowing almost everything about the reliability.

So, as we this data once we do the analysis and once we fit the data whatever parameters we get, that becomes the basic data which is used in all other matters which we will discuss like system reliability approach and all other methods which we have discussed earlier for that reliability, in this in this lecture series.

(Refer Slide Time: 01:43)

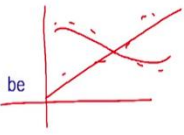


Least Square Estimation



NPTEL ONLINE CERTIFICATION COURSES
 INTRODUCTION TO RELIABILITY ENGINEERING


- Given a line equation $y = mx + c$ (fitting probability) 7.7P
- Given data (x_i, y_i) , the parameters of the equation can be estimated as
 - $\hat{m} = \frac{\sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$ (7.1)
 - $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ (7.2)
 - $\hat{c} = \bar{y} - m \bar{x}$
 - $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ (7.3)
- Coefficient of determination
 - $r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{c} - \hat{m} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - Where r is called index of fit.




2

Dr. Neeraj Kumar Goyal

Indian Institute of Technology Kharagpur

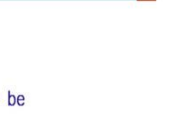


Least Square Estimation



NPTEL ONLINE CERTIFICATION COURSES
 INTRODUCTION TO RELIABILITY ENGINEERING

- Given a line equation $\hat{y} = mx + c$
- Given data (x_i, y_i) , the parameters of the equation can be estimated as
 - $\hat{m} = \frac{\sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$
 - $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
 - $\hat{c} = \bar{y} - m \bar{x}$
 - $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$
- Coefficient of determination
 - $r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{c} - \hat{m} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - Where r is called index of fit.



2

Dr. Neeraj Kumar Goyal

Indian Institute of Technology Kharagpur

We will be initially briefly discussed we because this is not a statistics class this is more about reliability. So, we are selectively discussing about the few principles which we are using for the reliability estimation. So, one way of doing that estimation is using the least squares estimation, where we try to take the square of error and we try to minimise that.

So, generally for least squares estimation when we are using we do the curve fitting. So, rather than curve fitting, as we know that if we want to fit let us see if we have the data. And if we want to fit some other curve, it would be difficult, but it is much easier if you have the data to fit the data to this straight line, because when we fit the data to a straight line, we can

see that the line should be passing to the reason where most of the points are lying nearer to the line.

Given a line equation

- $y = mx + c$

Given data (x_i, y_i) , the parameters of the equation can be estimated as

$$\begin{aligned}
 - \hat{m} &= \frac{\sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} \\
 - \bar{x} &= \frac{\sum_{i=1}^n x_i}{n} \\
 - \hat{c} &= \bar{y} - m \bar{x} \\
 - \bar{y} &= \frac{\sum_{i=1}^n y_i}{n}
 \end{aligned}$$

Coefficient of determination

- $r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{c} - \hat{m} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
- Where r is called index of fit.

So, this curve fitting is mostly done by the line fit line fitting. So, here we try to fit this line in a manner that our error in y is minimum or similarly if we reverse this equation, we can do the minimization of error on part of x , but here since like y here is the generally in all these cases y which we evaluate further and see is generally related with the failure probability and the x is mostly related with the other random variable which is time to failure.

So, we try to find out we try to transform our equation, the accumulative property distribution, and try to estimate try to get this data into the y equals mx plus c form. Once we are able to do that, then we estimate the value of m and c which represent the parameters of the distributions, which are related with the parameters of the distribution.

So, using those equations, we are able to evaluate m and c . So, we have the general equation whenever any data is available, let us say we have the x_1, y_1 and we have n such instant like x_1, y_1 that means, x_1 is that time is the independent variable, which is which can be time will some relation of time and y_1 is the dependent variable which is some relation of the probability of the failure. So, as such let us say we had n number of data or we have n number of failures. So, like we have done earlier, we have the n time to failure.

So, we have related with the time to failure we get the x_1 to x_n and then we have based on the ranking, we try to find out the F_{t_i} , F_{t_i} becomes F_{t_i} is related. Based on the F_{t_i} we calculate the y_1 to y_n . So, we have each pair, we have nine such pair. Using those pair, we can get the regression line we can fit and by using that we get the line after line fitting but is the values of m and c .

The value of m is statistically given as the estimated value of m for such case for n such data is given us summation of i equal to 1 to n $x_i y_i$ minus \bar{x} summation of y_i from i equal to 1 to n , this summation of i equal to 1 to $m y_i$ can also be written as n into \bar{y} , because \bar{y} is nothing but the 1 upon n summation of y_i .

So, that can also be written if we want we can use we can also replace these quantities as n into $\bar{x} \bar{y}$. Then, in denominator we have $\sum_{i=1}^n x_i^2$ minus n into \bar{x}^2 and we know what is \bar{x} , \bar{x} is the mean value of x_i that is the average value of x_i . So, 1 upon n summation of x_i .

Similarly, we can estimate the value of c the constant here, this value of c we can estimate by. So, value of c is nothing but if we see c is equal to $\bar{y} - m\bar{x}$, So, c is nothing but average value of y minus m into average value of x .

So, this value we calculate using this formula, so, m value of m will be the estimated value here, which we have estimated here. So, m estimated \bar{x} and \bar{y} is as you know that is nothing but the 1 upon n summation of y_i . So, based on this we are able to get the m and we are able to know the c and once we know the m and c based on the relationship which we develop we can find out the distribution parameters. Here as we know that we want to know how well this data is fitting to the distribution or the line straight line.


So, how well this data is fitting to the straight line if you want to find out we can use a coefficient of determination. So, coefficient of determination is 1 minus if you see this part this is y_i , y_i is the actual observation which you have and \hat{y}_i is the. So, this means y_i minus \hat{y}_i minus \hat{y}_i minus \hat{y}_i , same that is supposed to be 0 , if your points are lying exactly on the line, then y_i will be same as the \hat{y}_i or \hat{y}_i , \hat{y}_i is like when we put x_i value then based on the estimated value of m and c should give the nearest near value to the y_i .

So, this is kind of error and we are taking the square of the error as you see discussed earlier it is the least square estimation which is nothing but the error is square and we want to


minimise this. So, r square is giving this so, if this is exactly if it if your line is passing through all the points what will happen this value will be 0 and your r squared value will be 1, if it is far away from the line, the points are far away from the line but will happen this value will be higher. So, this value whether it is lesser or more because of this square both will be added.

So, whether the value is higher estimated value is higher estimated value is lower in both cases error will be high and error will be square and added here. Similarly, it is divided by the y_i minus \bar{y} square. So, here square root of this value when we take r square is generally called the coefficient of determination and the square root of this is we are calling as is an index of it. So, we want our r value to be close to the 1, if it is close to the 1 our fit is considered to be good.

(Refer Slide Time: 08:43)



Exponential Distribution



NPTEL ONLINE CERTIFICATION COURSES
 INTRODUCTION TO RELIABILITY ENGINEERING


- CDF is given as failure rate
 - $F(t) = 1 - e^{-\lambda t}$
- Converting this to straight line equation
 - $1 - F(t) = e^{-\lambda t}$
 - $\ln(1 - F(t)) = -\lambda t$
 - $\ln\left(\frac{1}{1 - F(t)}\right) = \lambda t$

$\ln(1 - F(t)) = -\lambda t$
 $\ln\left(\frac{1}{1 - F(t)}\right) = \lambda t$
 $y = mx + c$
- Using LSE, the parameter λ can be estimated as

$\hat{\lambda} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$
C=0

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

$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{m}x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$



3
Dr. Neeraj Kumar Goyal
Indian Institute of Technology Kharagpur



Least Square Estimation



- Given a line equation

$$y = mx + c$$

- Given data (x_i, y_i) , the parameters of the equation can be estimated as

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

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\hat{c} = \bar{y} - \hat{m} \bar{x}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

- Coefficient of determination

$$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{c} - \hat{m} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Where r is called index of fit.



Now, let us see that if we want to do this fitting of the data to the exponential distribution. So, if we find out that and we assume that our data is fitting to the exponential distribution, then what will be the exponential distribution parameter, we know that exponential distribution parameter is only 1 that is lambda, which is also called the failure rate because exponential distribution has the constant failure rate which is lambda. So, here for exponential distribution CDF cumulative distribution function or we can say the unreliability capital Ft is given as 1 minus e to the power minus lambda t.


So, now, this equation we have to change into the straight line equation. So, for changing into the straight line equation, we make some modification to this like what we can do, we can take 1 minus the side, so this will happen 1 minus Ft will be equal to e to the power minus lambda t. Now, again, what we can do we can take log of because this is exponential term, so we will take log, so log of 1 minus Ft will be equal to log of exponential minus lambda t will give you minus lambda t. Now, this is minus so I can write it that minus ln 1 minus Ft or I can write as 1 minus Ft 1 upon we know minus log of x is equal to log of 1 upon x. So, this will be equal to lambda t.

So, here now, if you see this equation is kind of turned into a linear equation, so, what linear equation this can be I can call it y, this lambda I can call this m and this t I can call it x. So, my higher c is 0, c is 0. So, my equation is turned into simple equation y equal to mx, since it is c equal to 0 that means, when I am plotting this line, the my lines should pass to the origin that is 0. So, for exponential distribution whenever we are fitting the intercept has to be 0, it

has to pass through the origin. So, now, we if we apply the formula which we have seen earlier.

So, y is equal to \ln of 1 upon 1 minus F_t and x is equal to t , once you put it here, then we will be able to. Now, what is m here m , as we saw this formula this is the value of m which we can calculate from here. Now, m value when we are calculating here we can calculate a summation of i equal to 1 $x_i y_i$ divided by summation of i equal to 1 x_i square. So, this value once we calculate this will give us the λ value. And y_i is this x_i is this and r square value again we can calculate, here c is 0 because c is 0 we do not need to calculate c here and r square value is 1 minus whatever formula is the same formula by putting c equal to 0 we are able to get it. Once we use this, we will be able to get the fit the fit it and we will be able to get the distribution.

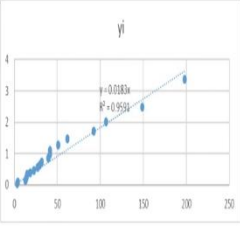
(Refer Slide Time: 12:04)




NPTEL ONLINE CERTIFICATION COURSES
INTRODUCTION TO RELIABILITY ENGINEERING

4

i	t_i	$F(t_i)$	x_i	y_i	$x_i y_i$	x_i^2	$\frac{[(y_i - m x_i)]^2}{m^2 x_i^2}$	$\frac{[(y_i - y)]^2}{n^2}$
1	3.3	0.034314	3.3	0.03492	0.115224	10.89	0.000652	0.850229
2	4.2	0.083333	4.2	0.08701	0.365448	17.64	0.000101	0.756872
3	12.9	0.132353	12.9	0.14197	1.831416	166.41	0.008903	0.664265
4	13.8	0.181373	13.8	0.20013	2.761741	190.44	0.002776	0.57285
5	14.3	0.230392	14.3	0.26187	3.744801	204.49	9.96E-09	0.483193
6	14.8	0.279412	14.8	0.32769	4.849774	219.04	0.003198	0.396028
7	18.5	0.328431	18.5	0.39814	7.365573	342.25	0.003507	0.31232
8	22.8	0.377451	22.8	0.47393	10.80567	519.84	0.003163	0.233349
9	27.1	0.426471	27.1	0.55595	15.06614	734.41	0.003538	0.16084
10	29.7	0.47549	29.7	0.64529	19.16515	882.09	0.01024	0.097159
11	32	0.52451	32	0.74341	23.78909	1024	0.024704	0.045619
12	39.5	0.573529	39.5	0.85221	33.66237	1560.25	0.016532	0.01098
13	41.3	0.622549	41.3	0.97431	40.23919	1705.69	0.047395	0.0003
14	41.6	0.671569	41.6	1.11343	46.31858	1730.56	0.123427	0.024471
15	51.1	0.720588	51.1	1.27507	65.15601	2611.21	0.114869	0.101171
16	61.7	0.769608	61.7	1.46797	90.5739	3806.89	0.113999	0.261098
17	92.2	0.818627	92.2	1.7072	157.404	8500.84	0.000328	0.562811
18	106.6	0.867647	107	2.02228	215.5754	11363.56	0.004814	1.134839
19	148.8	0.916667	149	2.48491	369.7541	22141.44	0.058124	2.334514
20	198.1	0.965686	198	3.37221	668.0348	39243.61	0.066025	5.833263
				0.957	1778.578	98975.55	0.806298	14.83817
			λ		0.01832			
			R^2		0.959134			
			R		0.979354			





So, here I have taken one example here, this example I will try to show you in excel sheet I have developed this in Excel sheet I will try to show you. So, that you can follow up and you can see that how this is done.

(Refer Slide Time: 12:33)

Distribution Fitting Excel

#	TI	Sturges Rule	$\sqrt[3]{1+3.3n}$	Interval	t_c	Nf	Ns	Rs	F(t)	$f(t_j)$	$z(t_j)$			
3	13	20	20	0	175	0	35	1	0	0.5143	0.5143			
4	15	31	2025	337.5	350	525	18	17	0.486	0.514	0.5882			
5	10	36			700	875	10	7	0.2	0.8	0.0857	0.4286		
6	16	47	Dividing the data in 6 intervals				1050	1225	3	4	0.114	0.886	0.0286	0.25
7	3	98			1400	1575	1	3	0.086	0.914	0.0571	0.6667		
8	5	157			1750	1925	2	1	0.029	0.971	0.0286	1		
9	6	182			2100	1925	1	0	0	1				

Distribution Fitting Excel

#	TI	Sturges Rule	6.09542	Interval	t_c	Nf	Ns	Rs	F(t)	$f(t_j)$	$z(t_j)$			
3	13	20	20	0	175	0	35	1	0	0.5143	0.5143			
4	15	31	2025	337.5	350	525	18	17	0.486	0.514	0.5882			
5	10	36			700	875	10	7	0.2	0.8	0.0857	0.4286		
6	16	47	Dividing the data in 6 intervals				1050	1225	3	4	0.114	0.886	0.0286	0.25
7	3	98			1400	1575	1	3	0.086	0.914	0.0571	0.6667		
8	5	157			1750	1925	2	1	0.029	0.971	0.0286	1		
9	6	182			2100	1925	1	0	0	1				

Distribution Fitting Excel

7	3	98			1400	1575	1	3	0.086	0.914	0.0571	0.6667
8	5	157			1750	1925	2	1	0.029	0.971	0.0286	1
9	6	182			2100	1925	1	0	0	1		

So, let me show you one Excel sheet here, I have developed this Excel sheet for your purpose and you can follow it up. So, let us go with exponential i before exponential distribution sometimes we may be interested in histogram plotting also. So, like we discussed in earlier classes that if you have the time to fill the data, which you may have it here you can arrange the data in increasing order then you can divide the data into interval by using the Sturges formula.

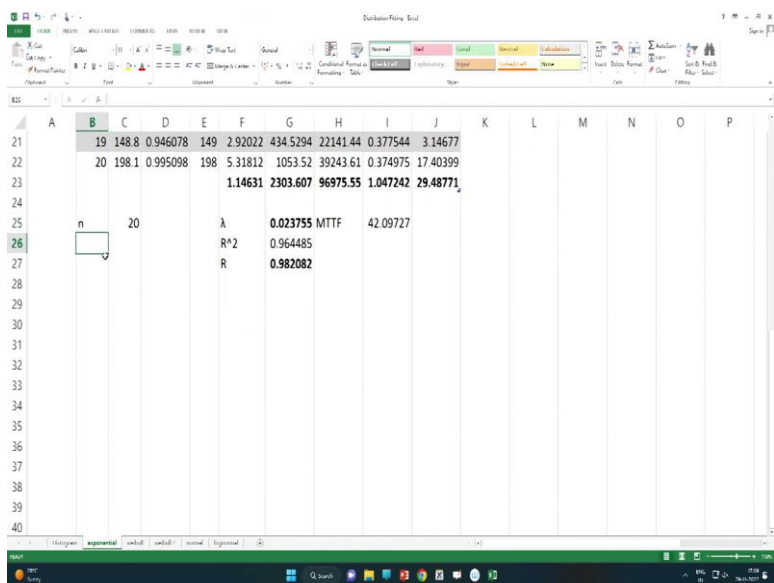
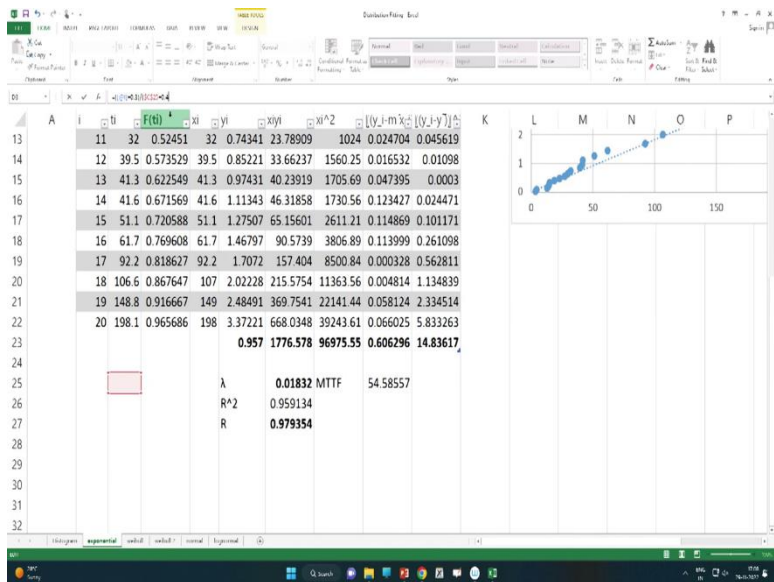
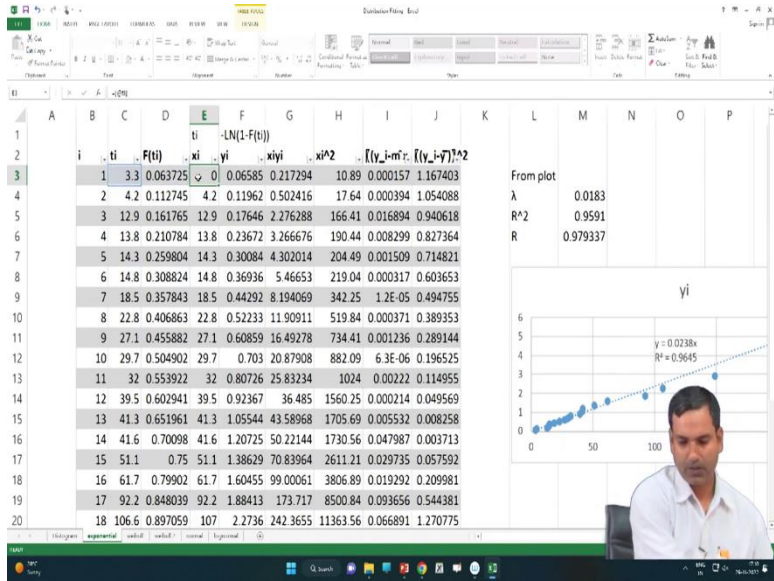
So, Sturges Rule when you apply that is $1.3 + 1.3 + 1.1 + 3.3$ into $\log_{10} m$, m is 35 here. So, when you use that value, you get this value as 6 around. So, for 6 intervals, you can divide data into the 6 interval and for the 6 interval you will get the number of failures you will get the number of survival. So, initially 35 units of working.

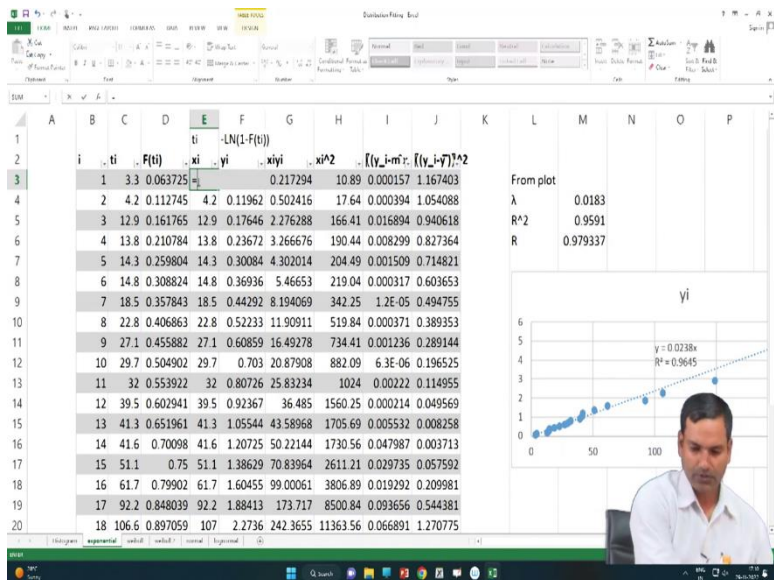
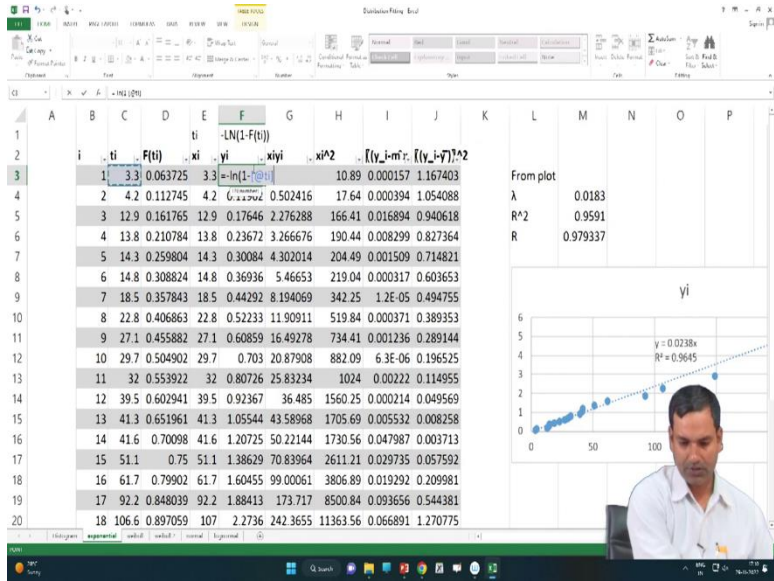
So, if you see that here up to first interval 350 18 units failed remaining 17 like the same way as we did earlier we can do this analysis and we are able to get the ft small ft. So, this is the analysis using the nonparametric approach, which we have already studied in previous classes.

Now, here sometimes we can plot this frequency curve or histogram curve to see that how this ft is varying or how this number of failures are changing with time. So, this can give us sometimes an idea or better will be if we plot the failure rate curve right. So, zt if you plot the zt curve here that can give us a little bit idea whether failure rate is increasing or decreasing constant like that.

So, that can indicate us whether we should use exponential distribution or should we use variable distribution etcetera. But we can also directly do by directly fitting the distributions and we are able to know which distribution is fitting.

(Refer Slide Time: 14:23)





Exponential Distribution

- CDF is given as

$$F(t) = 1 - e^{-\lambda t}$$
- Converting this to straight line equation

$$1 - F(t) = e^{-\lambda t}$$

$$\ln(1 - F(t)) = -\lambda t$$

$$\ln\left(\frac{1}{1 - F(t)}\right) = \lambda t$$
- Using LSE, the parameter λ can be estimated as

$$\hat{\lambda} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

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

$$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{\lambda} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Dr. Binay Kumar Goyal

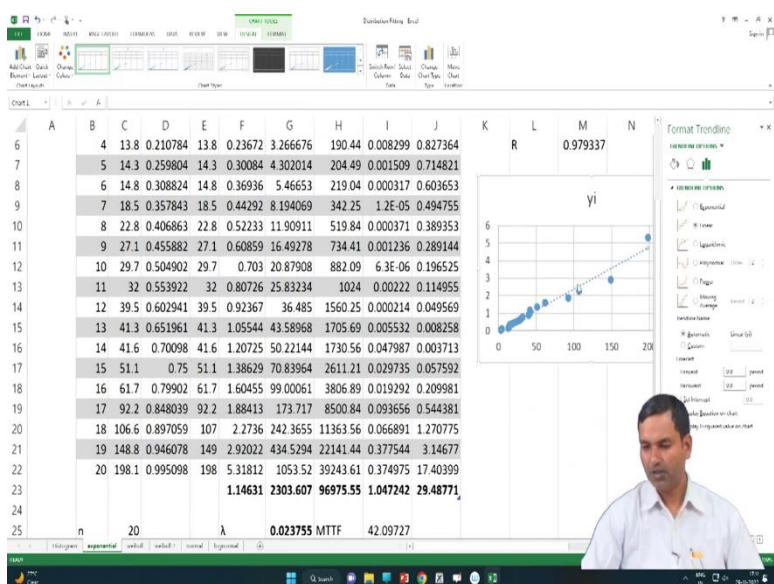
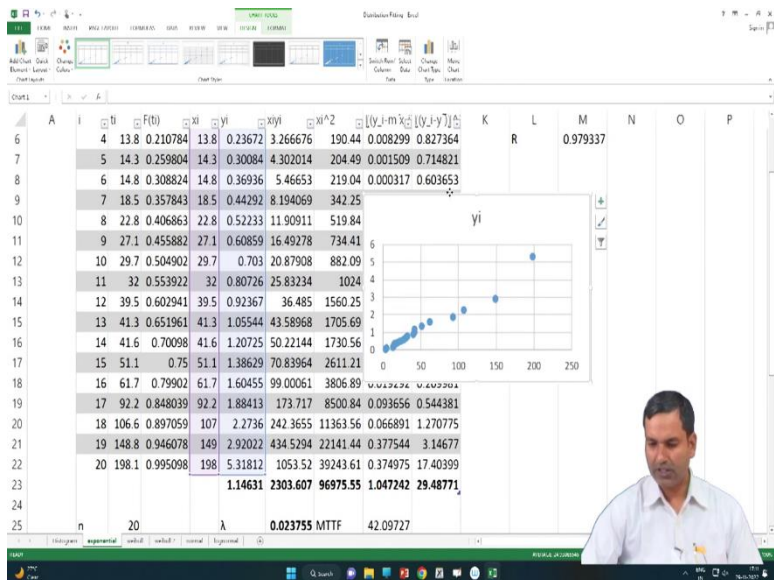
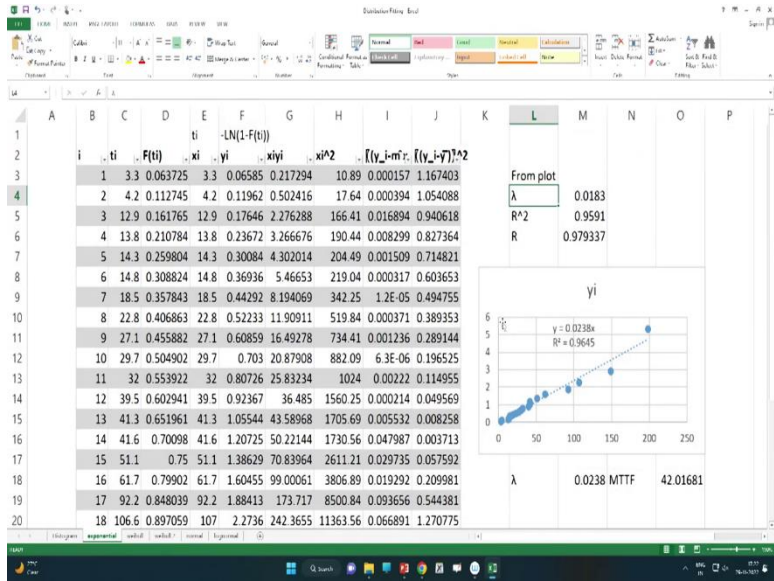
So, two distribution fitting let us say, I have these data's which I have already sorted. So, we have 20 data points here. So, what I did I took the data points then I sorted the data and this becomes the increasing order. Once I have then we can use the rank methods like we if we use the, as we know that we can use the median rank method.

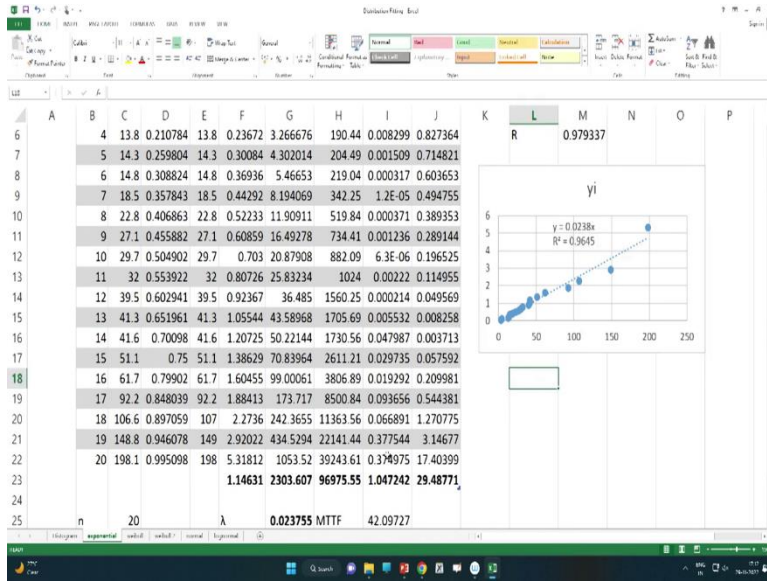
If you use the median rank method, then we know for median, the median rank is $i - 0.3$ divided by $n + 0.4$. So, same thing I will do this is equal to i , my value of i is in column B plus 0.3 divided by my value for n I have written below, let us say I will write it here somewhere. So I am using this value. And I am using this value, since I do not want this value to change, I will put dollar sign here. And this plus 0.4 .

So, $i - 0.3$ divided by $n + 0.4$, I am using the same here. And since I have not put I will put 20 here that is my value of n C25 right. So, we are able to have these values F_{t_i} . But as we know that F_{t_i} is not having the my axis t_i and y is F_{t_i} in general, I want to know how F_{t_i} is changing with the time, but that distribution is not following the straight line. So, I have put this straight line equation. So, for a straight line as we discussed earlier, but we have to know for straight line, we have to change F_t value into this \ln of $1 - F_t$ or \ln of $1 - F_t$.

So, same thing we have done y_i value when we get that is \ln of $1 - F_{t_i}$, that is equal to, I can take both either \ln of $1 - F_t$ or \ln of $1 - F_t$ to both will give the same. So, I will use this \ln of $1 - F_t$, F_t is this value. So this way, I will get the y value. Same values I will get, I have already done this. Now, what is x_i value here? Let us see what is the x_i here? Here x value is the t . So, whatever my time is there exactly same time I am supposed to take. So, this will be equal to x will be equal to t . So same value will appear here.

(Refer Slide Time: 17:51)

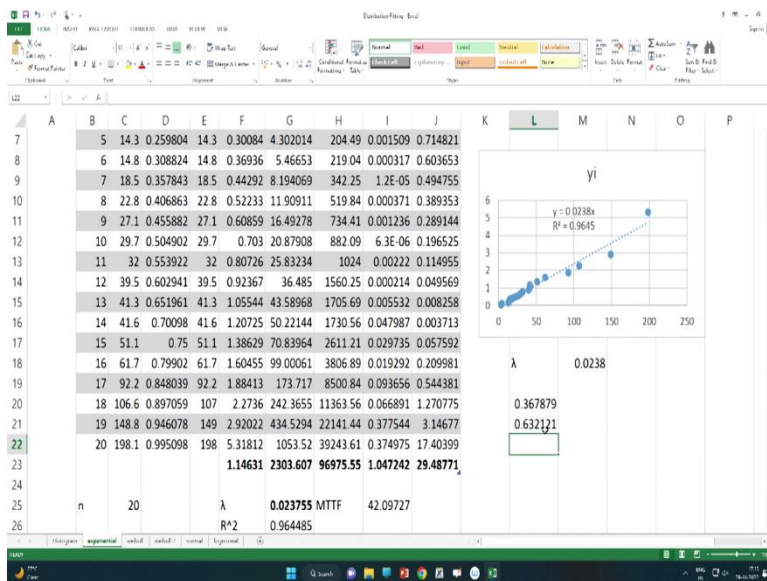




Exponential Distribution

- CDF is given as
 - $F(t) = 1 - e^{-\lambda t}$
- Converting this to straight line equation
 - $1 - F(t) = e^{-\lambda t}$
 - $\ln(1 - F(t)) = -\lambda t$
 - $\ln\left(\frac{1}{1 - F(t)}\right) = \lambda t$
- Using LSE, the parameter λ can be estimated as
 - $\hat{\lambda} = \frac{\sum_{i=1}^n t_i y_i}{\sum_{i=1}^n t_i^2}$
 - Where, $y_i = \ln\left(\frac{1}{1 - F(t_i)}\right)$; and $x_i = t_i$
 - $r^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{\lambda} x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$

Dr. Hitesh Kumar Goyal
Indian Institute of Technology Khargpur



Excel spreadsheet showing a data table with columns A through J and a scatter plot of y_i vs x_i . The plot includes a linear regression line with the equation $y = 0.0238x$ and $R^2 = 0.9645$. The regression statistics are:

- $\lambda = 0.0183$
- $R^2 = 0.9591$
- $R = 0.979337$

Additional values shown are $\lambda = 0.0238$ MTTF and 0.367879 .

Excel spreadsheet showing a data table with columns A through J and a scatter plot of y_i vs x_i . The plot includes a linear regression line with the equation $y = 0.0238x$ and $R^2 = 0.9645$. The regression statistics are:

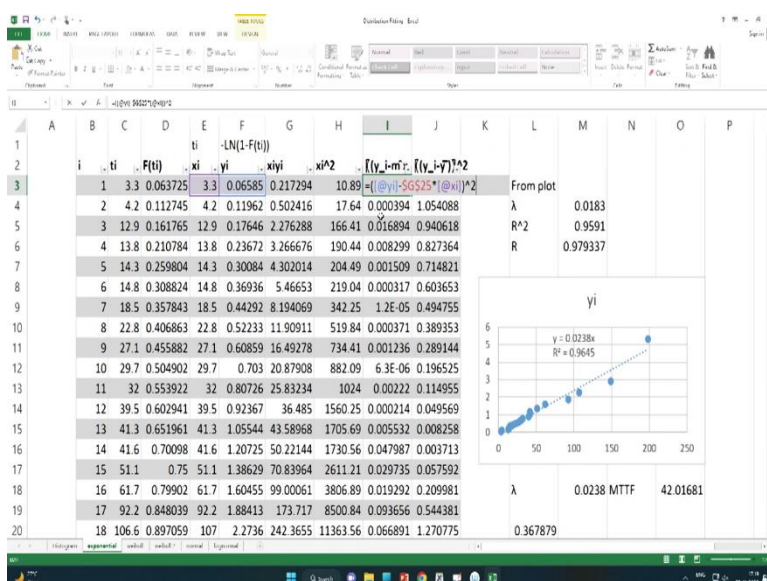
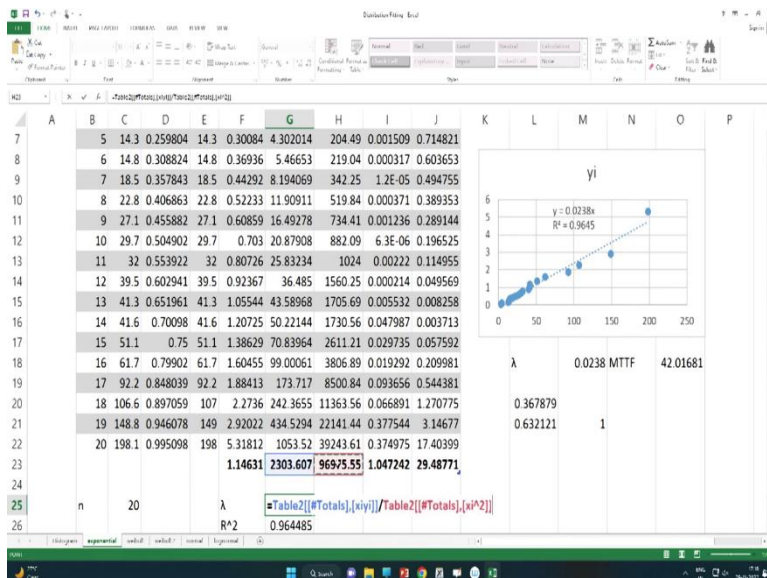
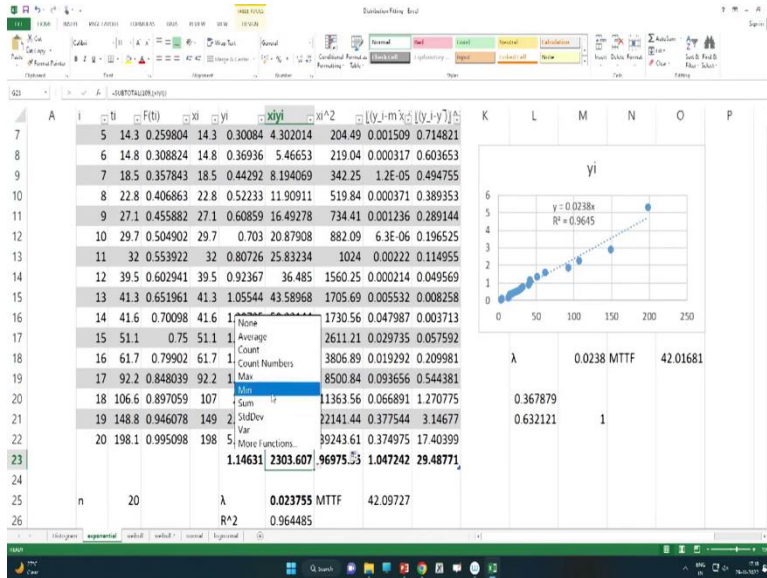
- $\lambda = 0.0183$
- $R^2 = 0.9591$
- $R = 0.979337$

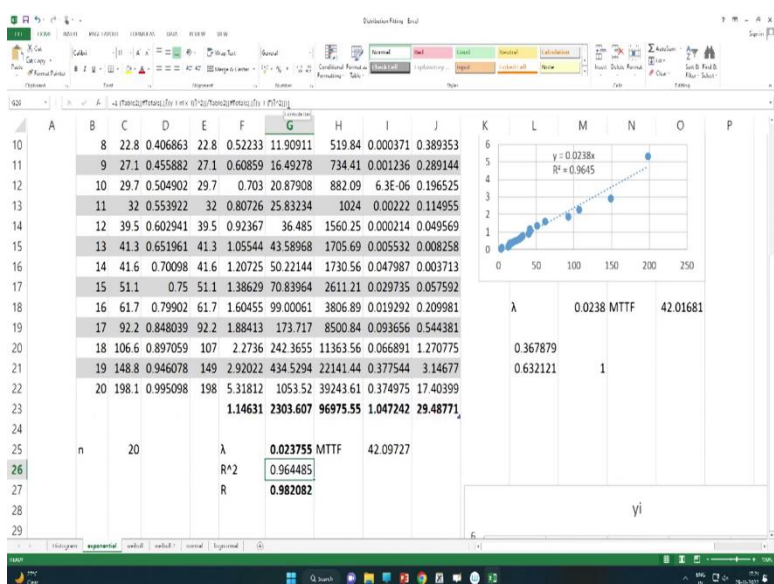
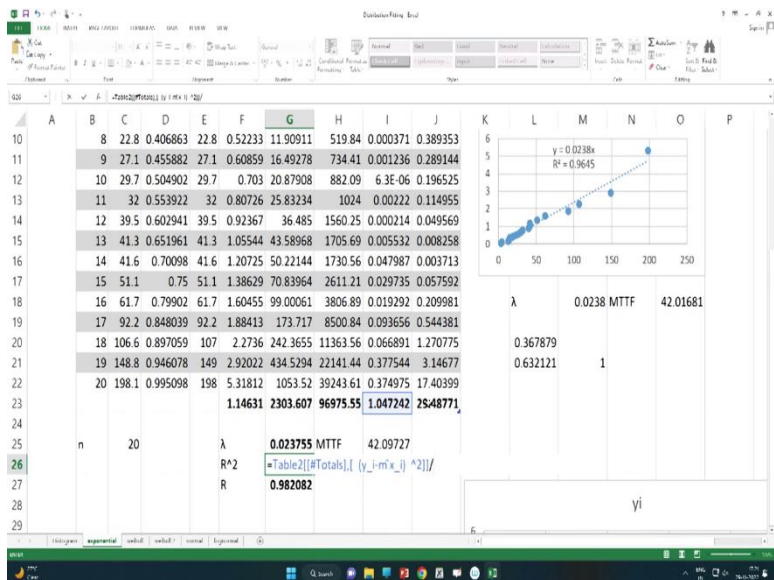
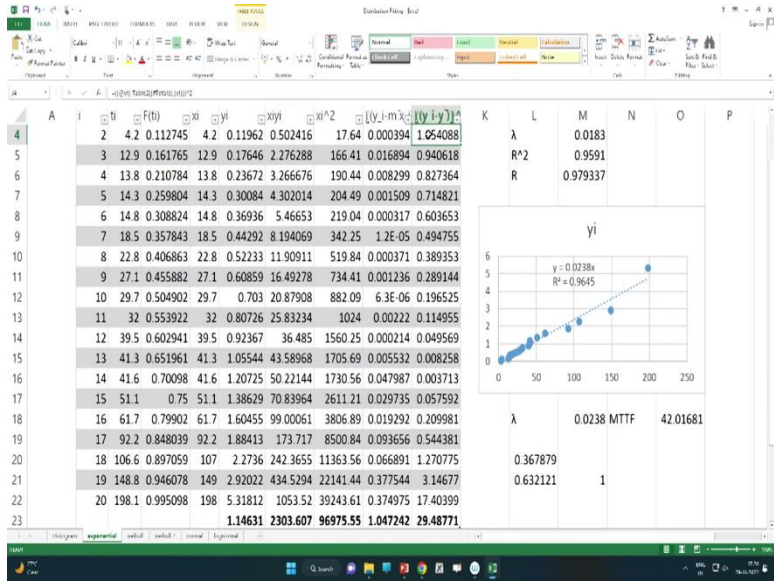
Additional values shown are $\lambda = 0.0238$ MTTF and 42.01681 .

Excel spreadsheet showing a data table with columns A through J and a scatter plot of y_i vs x_i . The plot includes a linear regression line with the equation $y = 0.0238x$ and $R^2 = 0.9645$. The regression statistics are:

- $\lambda = 0.0183$
- $R^2 = 0.9591$
- $R = 0.979337$

Additional values shown are $\lambda = 0.0238$ MTTF and 42.01681 .





Now, what I need to do is I, there are two ways, I will show you the graphical way, which we can use in the Excel. And I will also show you that how we can use the formula to calculate the same thing. So, let us first see the graphical way. So, I have already plotted here, I will do this exercise again in front of you so that you can also follow it. So, I will take this graph below, I will not use the same graph. I will make a new graph which, so what I have to do, I have to plot x_i vs y_i .

So same thing, I will take it here. And I will go to this insert, you can do use any statistical package also for this, but that is see that step by step, how we can do it. So I will use this dot xy plot scatter plot. So, I have done this xy scatter plot. Now, this xy scatter plot shows my y versus x_i . But, the problem is that I want to know, I want to fit this straight line to this. So, either I can take the scale manually and fit it here like this. But that would be little errors, but still I can do, or I will use the calculated way. So Excel has the inbuilt function there, it uses the trend line function.

So, we can use the trend line function here. If you see trend line, I get the various options here. So first option, as you know, for since we are fitting the log, we are fitting the exponential distribution, we have to set the intercept. Because we want intercept to be 0, we do not want that it should not have any intercept on x or y axis. So that should be 0. And whatever is the fitting line, I want to know the equation so that I can use the values. So, I will display the equation on the chart, I will also display the r square value on two charts so that I can determine how much is the index of it.

If you see by doing this, I am able to get this, I am seeing that y is equal to $0.0238x$. So, as we know from this, what is my y what is my m, m is lambda. So, from this my lambda which I have, which I want to calculate here, my estimated value of lambda is directly I am able to get it from this graph and how much is that this lambda value is equal to 0.0238.

Now, if I had know the lambda, I know that distribution. Now, whatever I want to plot, this is an estimated value of lambda from this plotting and based on it and index of it is also good it is almost r square is 0.96, this is like graphical way using the Excel you can do it easily, or you can also there are exponential graph paper are also available.

In exponential graph paper what happened they have modified the axis like this will be the logarithmic axis and this will be the linear axis because t is linear and $1 - F_t$ is the log the logarithmic axis. So, this will be the logarithmic axis and this will be the linear axis. So,

what will happen the same thing but what we are doing by the modification of data same thing can be achieved by the modification of axis.

So, that means 1 then it will appear 10 like that. So, because of that what will happen once you have that data you can do the plotting here, that will also look into this straight line, that will also supposed to follow on the straight line. So, same like exponential we have the graph paper for v bull etcetera, but we can use them we can plot the data on them and we can find out, there is another way to find the lambda here that is if we are able to know we know that when t is equal to lambda, if you see here, if t is equal to lambda here if you put t equal to 1 upon lambda here, then what will happen this will become e to the power minus lambda divided by lambda that will be e to the power minus 1. So, and e to the power minus 1 is how much that is equal to exponential minus 1.

So, whenever value of t is equal to lambda my reliability is 0.36 and unreliability is equal to. So, my CDF value is always this 62 point 63.21. Now, according to this if I want to calculate y value y value as we know is minus ln of 1 minus this value, minus ln 1 minus Ft. So, once we use this, so, here if you see our if I look at this curve, then my value would be falling somewhere here, for one value if I plot and if I take this intercept on x that would be somewhere around 40 here 40 to 45 like that. So, I can calculate that.

So, as we see the same value like point what we have is lambda is this 0.02. So, my MTTF value will be 1 upon lambda that will be equal to 1 divided by this value which is around 42. So, this intercept for 1 is around 42 time is equal to 42 that means, my mean time to failure is 42. Because they know that when t is equal to MTTF, then when I put t equal to MTTF, MTTF is 1 upon lambda Ft becomes 0.632 and my if I take the 1 minus ln of 1 minus Ft that becomes to be the 1. So, by plotting from the plot itself, I will also I can find out what is my lambda or what is my MTTF.

Once I know MTTF from here I can do but for that I need to have this graph properly marked and I should also have the minor axis properly marked. So, that I can observe clearly how much correctly it is obtained. So we are able to get lambda we are able to get MTTF. Now, let us see the another method the method this is also using LSE but here is the visual way of doing that type we can use Excel sheet graph approach to do that.

Now let us see we want to calculate the same thing. So what I have done here, this approach I will be using for across the distributions. So I have used the generalised approach rather than

using this formula here, this formula also I have used but let us say if I do not use this formula, let us say we use the same formula. So for this we need to calculate two values $x_i y_i$ and x_i^2 . So, first let us see how do we do that so we calculate $x_i y_i$. What is $x_i y_i$? $x_i y_i$ is x_i multiply by y_i , we get this $x_i y_i$. Then another value which we need is the x_i^2 . So, that is equal to x_i^2 . We got this value x_i^2 also.

Now, using these what we have to do we need to take the summation of this. So, this is the summation I have got, I have used the sum of this value, I have used the sum of these values. So, this is sum of $x_i y_i$, this is sum of x_i^2 . Now, what is my formula here? My formula of λ is $\sum x_i y_i$ divided by summation of x_i^2 . So, λ is equal to summation of $x_i y_i$, that is this summation of $x_i y_i$ divided by summation of x_i^2 and this becomes my λ . If you see this λ and this λ both are same.

Whatever we have done by calculation, same thing has been done by the excel in drawing or plotting this trend line, we get the same value. And if I want to calculate r^2 , what is my r^2 , r^2 is summation of $i=1$ to n $(y_i - m x_i)^2$ minus $m^2 \sum x_i^2$. So, here I am need to take $y_i - m x_i$. So, how can we can I take $y_i - m x_i$ that is equal to y_i , y_i is how much? y_i is this one, minus m value I have already calculate this λ is my m value.

Now, this value I do not want to change, so I will put the dollar here again. So that my values do not change, m multiply by x_i . So, this gives me the now these values have to take this square, if you remember the formula, $(y_i - m x_i)^2$, I have already done $y_i - m x_i$. Now, let us do the square. So, for square I will do the squaring of this. So, once we do this and then again I can take the sum of this, the sum will be this value, which we are having here.

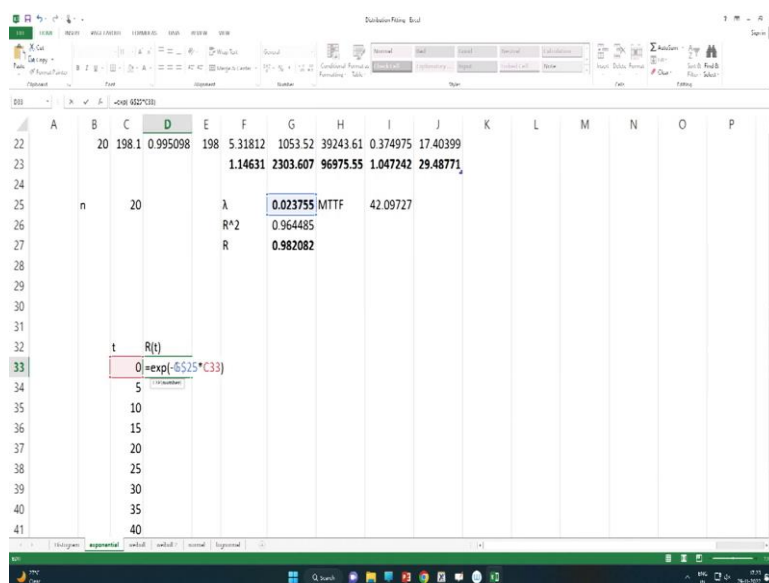
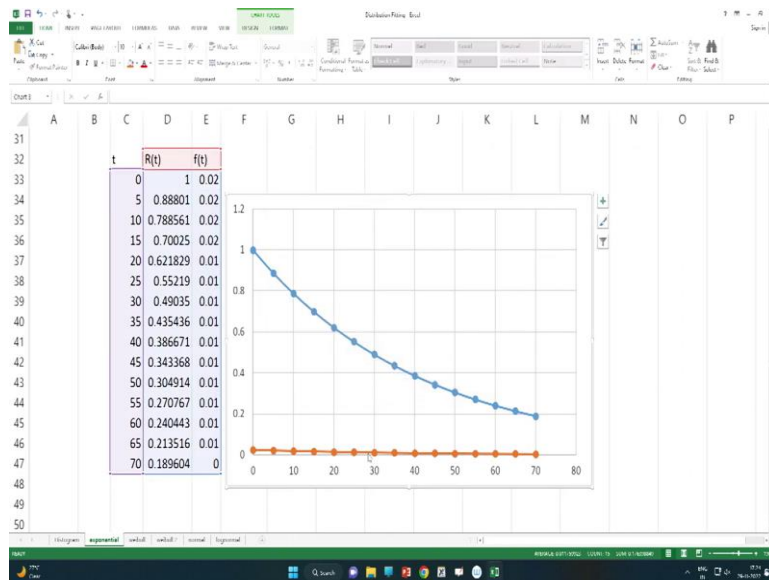
So, by taking the sum, I will get this value. Now, let us take the denominator is $\sum (y_i - \bar{y})^2$. So $\sum (y_i - \bar{y})^2$, that means I will take this, this is equal to $\sum y_i^2 - 2 \bar{y} \sum y_i + n \bar{y}^2$. So what I have done, I have taken the average value already here, $\sum y_i^2 - 2 \bar{y} \sum y_i + n \bar{y}^2$. $\sum y_i^2 - 2 \bar{y} \sum y_i + n \bar{y}^2$, this is average value, I will show you that this value is nothing but the average value, this is the I have taken the average value here.

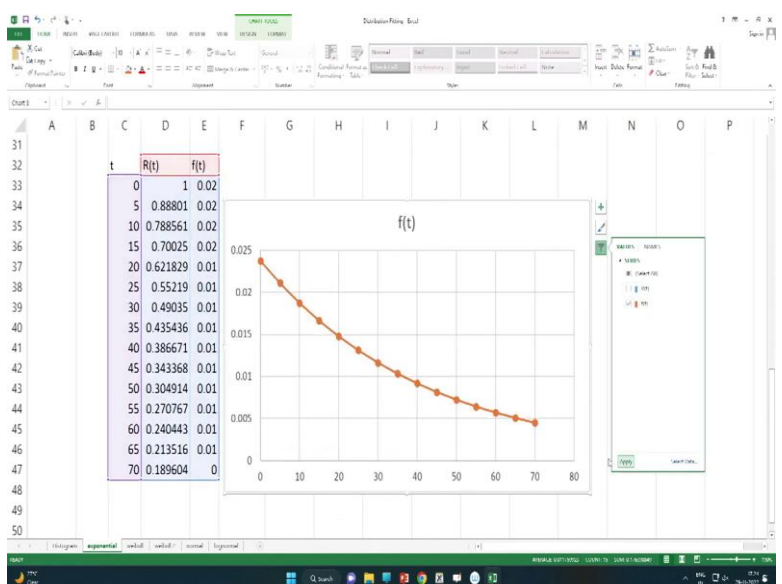
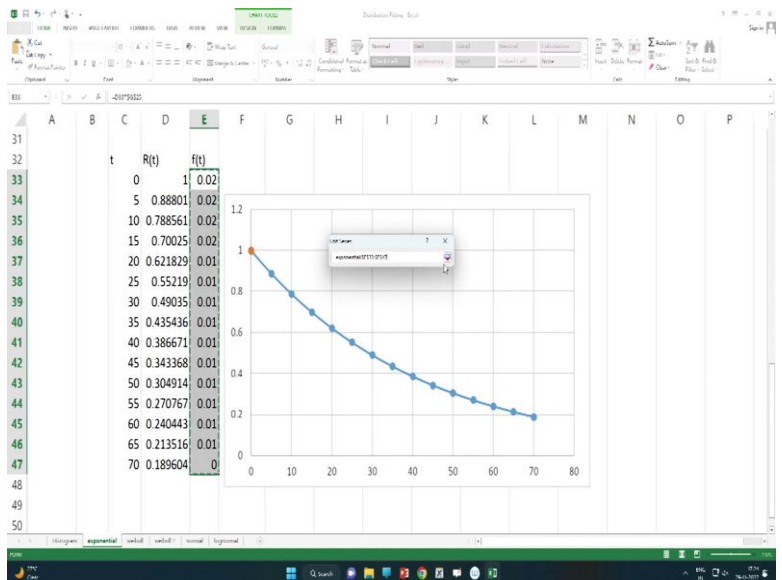
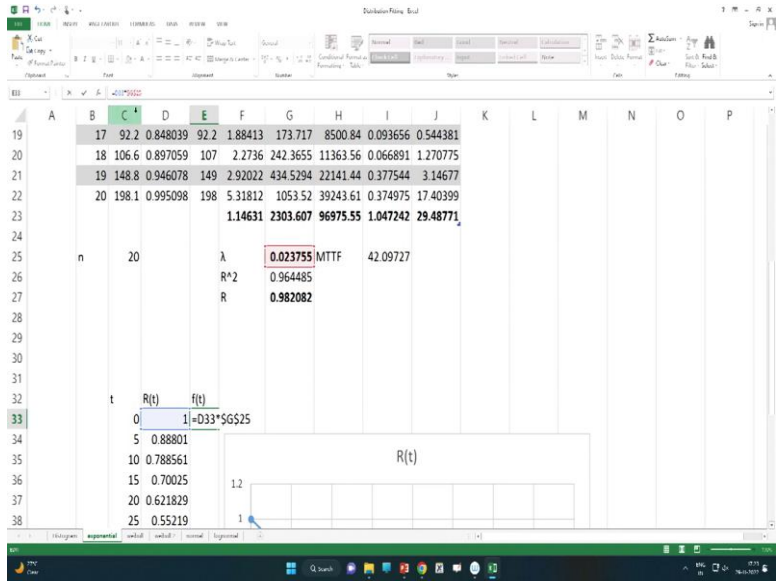
So, $\sum (y_i - \bar{y})^2$, this is $\sum y_i^2 - 2 \bar{y} \sum y_i + n \bar{y}^2$. So, I am able to get the same formula, this is my denominator. So, this summation, which I have got divided by this will give me the r^2

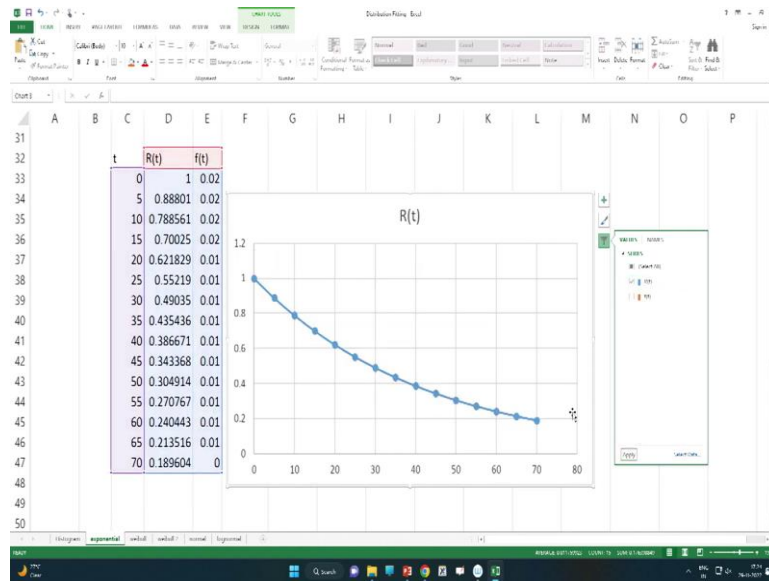
value. So r square value is nothing but this divided by this sorry, and I have to subtract the whole value from the 1, 1 minus of this, 1 minus of this because I want this value to be near to the 0 r square is 0.964485, which is also if you see this is also same as what we have got here. And r value is the square root of this.

So, my index of fit is 0.982, which is a very good fit. And once this value I am removing, because we have already done this exercise. So, we are able to get lambda we are able to get r square we are able to get r everything we are able to get. And once we have this we are able to fit the exponential distribution. And if you want to know we, once know the per parameter of the distribution, we know everything about the distribution.

(Refer Slide Time: 29:06)







So if let us say if I want to take some time values, and if I want to know the R_t , how can I do is let us take I will take some time 0, let us say I will take the difference of 5, 0, 5, let us take some values. Some more values we can take an R_t is we know R_t is equal to exponential minus lambda t and how much is lambda? This is my lambda multiply by t values given here.

Now this G25. As I discussed, this need to be near the constant. This should not change. So, I will use the dollar sign here. And If I want to plot R_t , I can plot R_t . I will just go to the insert. And I will plot the xy chart with line. If you see this is my exponential distribution, which I have fitted and if I want to know the Z_t , Z_t is constant line. So, no need to do that if I want to know F_t , what is F_t ? $F_t = Z_t R_t$. So, that is equal to R_t into lambda t, my lambda t is this value, because this is constant. So, I will use again dollar sign here and this becomes my F_t .

I can add the F_t value here; I can select data I can add one curve here, series name is this and the x values are already taken here, and y values are we have to take from F_t and we get this. You see, this is my R_t curve and this is my F_t curve because valued are different so I will plot one at a time, this is my F_t and this is my R_t . As you see that this is exponentially decreasing and whenever I want to know the value, I can know the value.

So, we stop it here today, we will continue this discussing this more distribution how they can be fitted to the data in next classes.