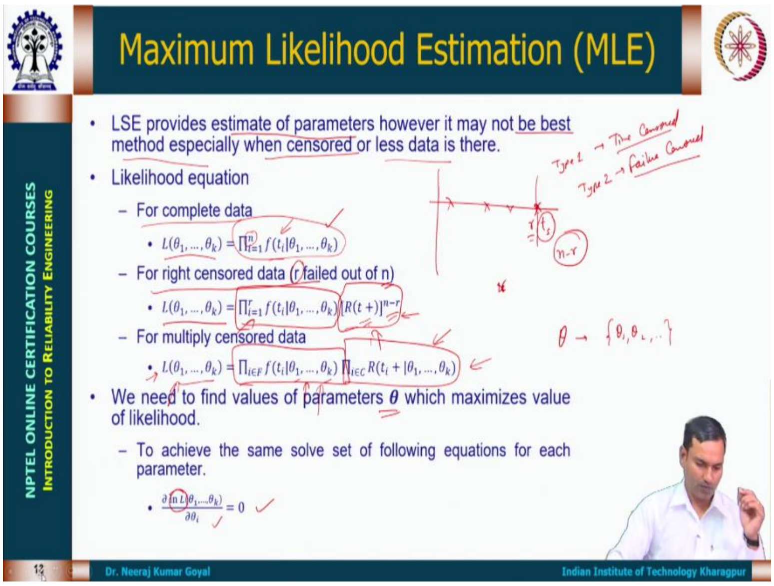
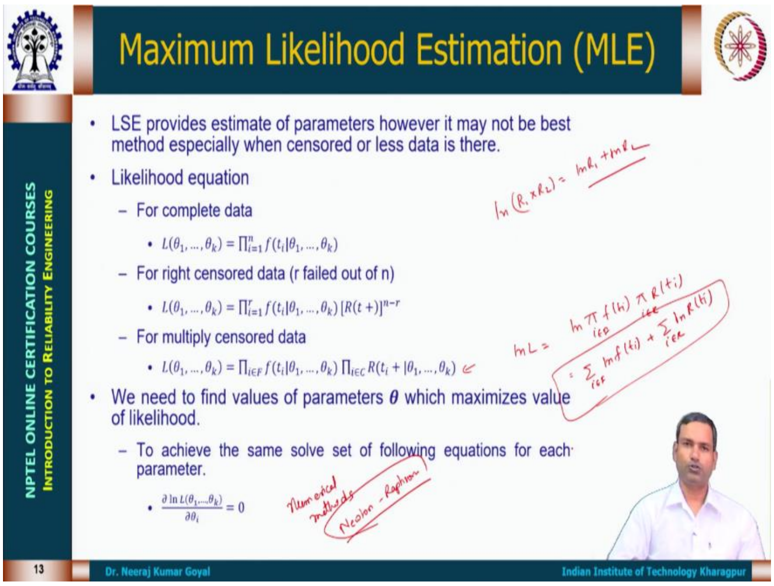
Introduction to Reliability Engineering
Professor Neeraj Kumar Goyal
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur
Lecture: 32
Failure Data Analysis (Parametric) (Contd.)

(Refer Slide Time: 00:26)

Hello everyone, we have been discussing about how to fit distributions to the failure data that means the failure data time to failure data or that may be the repair data time to repair data. So, once we have time to repair data or once we have the time to failure data, we are able to know the liability, we are able to know maintainability and if you know that two, we are able to get an understanding about the availability also. So, here will continue our discussion. So, last time we discussed about the least squares estimation where we try to fit line which is giving us the least error after converting the equation into the straight line equation, the now we today we will discuss about maximum likelihood estimation.

(Refer Slide Time: 01:17)





So, as we discussed LSE or least square error square estimation provides an estimate of parameters. So, we are able to do that, but essentially it is not the best method because here we are doing lot of convergence here like we have to first convert into the straight line equation then we have to see which one is fitting and especially, so, this may be giving a reasonable estimate when data is large and complete, but when sensor data is up there or when data sample size is small. In that case, LSE may not be able to provide a good estimate here means that estimation error would tend to be large that whatever is the true value for the process, it may have little larger error.

For complete data

- $L(\theta_1, \dots, \theta_k) = \prod_{i=1}^{n} f(t_i \mid \theta_1, \dots, \theta_k)$
  For right censored data (r failed out of n)
- $L(\theta_1, \dots, \theta_k) = \prod_{i=1}^{r} f(t_i \mid \theta_1, \dots, \theta_k)[R(t+)]^{n-r}$
  For multiply censored data
- $L(\theta_1, \dots, \theta_k) = \prod_{i \in F} f(t_i \mid \theta_1, \dots, \theta_k) \prod_{i \in C} R(t_i + \mid \theta_1, \dots, \theta_k)$

So, MLE is considered to be a better method for making an estimation. However, MLE requires a lot of mathematical manipulation and computations as we will see. So, MLE as it says it is the maximisation of likelihood. So, by maximising the likelihood we are trying to find out the parameter. So, for parameter estimation this should be primarily used if possible. So, here as we see that we have a likelihood equation here likelihood here for complete data means all failure data if you have so, this is nothing but a if you see this kind of joint probability density function of all the times to failure data which you have got or it may be the time to repair data.

So, if you have n number of devices, which are put on test and all n devices are failing at time t1, t2, tn, so, you will be having a combined distribution of that. So, generally we assume that the failure times are independent to each other, since, we are assuming independency So, f ti once we multiply the PDF for each time to failure, we get the combined PDF for whole observation set and that makes our likelihood generally, so, this is for failure data. So, generally for failure data we use this product density function, but if you have the censoring, in case of censoring let us say we put n units on tests r failed means n minus r units are still working. Right censored means as we discussed here, there are two types of sensor data in right sensor type 1, type 2, type 1 is time censored and type 2 is failure censored as you discussed earlier, then let us say we have a test time is fixed ts let us say.

So, whenever fail whatever the number of failure happens that will be recorded, but test will stop here. But in failure censored if we have decided k number of failures or r number of failures in this case, then rth failure will be here. And when rth failure happens then test will stop. So, it makes very little difference. Because whatever is the time at which you are censoring, at that time what is happening the devices r devices are failed. So n minus devices are still working. So that means n minus r devices are reliable. So, here in likelihood equation in case of censoring the sensor data which we have for the sensor data, the user reliability equation and for failed data we use the PDF equation. So, for failed data we get the PDF value add that value and multiply.

So, for each failure point we will have the corresponding PDF value we will multiply and for each value for which the reliability that is the censoring is happening for right censored data all points are censored at same place. So, n minus r unit have been since censored let us say t plus time or we can say ts time whatever we try to denote. So, reliability at that time will be multiplied for all censored values. So, that will become n power n minus r. So, r units goes here and n minus r units goes here similar is for multiply sensor data same thing multiply sensor data we can say the generalised point that for all the sensors unit the reliability is supposed to be multiplied and for all the failed units, we will multiply the probability density function and that makes our likelihood.

Now, we want to get the value of theta1, theta2, theta3 these are the set of the parameters for the distribution like for exponential distribution we have only one parameter so, we will have theta 1 as the lambda for variable distribution two parameter variable distribution we have two parameters. So, let us say theta 1 is equal to theta and theta 2 can be the beta. So, here these are the set of parameters used by the distribution for normal distribution may be mu and sigma for log normal it can be T median and another parameter can be as so, that way we will be, so, this likelihood equation are the function of these parameters because everything else is supplied t1 is known.

So, all t1, t2, t3 is known. So, that is supplied because that is the observation points only thing which is unknown in this equation is the parameters of the distribution and we want to find out the values of these parameters at which this likelihood value is the highest value. So, we want to maximise likelihood for a given set of parameters here. So, here theta is a set of like capital theta which you are writing it is the set of parameters this theta 1 theta 2 or whatever is the parameters involved. So, how do we maximise any function to maximise any function we can take the derivative of the function and equate it to the 0 because at maximum value we expect that slope will be 0, but maximisation doing at the directly at the likelihood function is not advisable here because that involves lots of multiplications here as we see here.

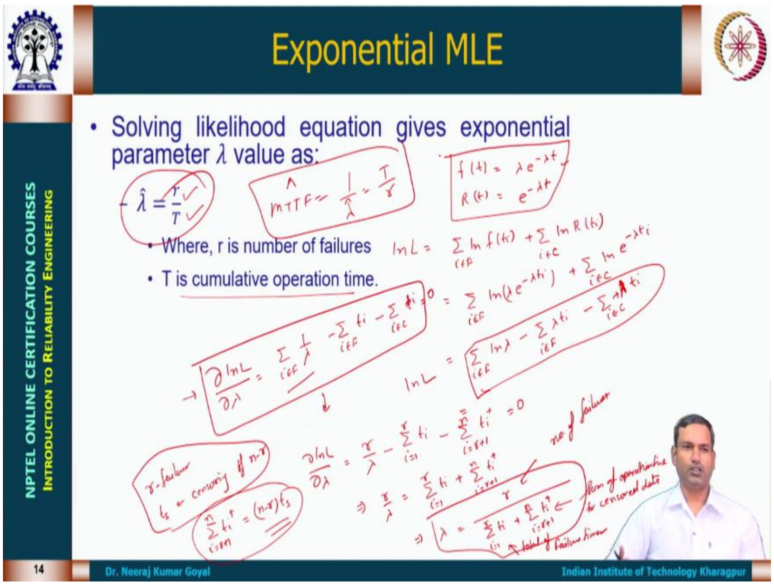
So, we use the property that wherever the likelihood is maximum the log of likelihood is also maximum. So, log of likelihood has the same point of maximum or same place where the likelihood will be maximum. So, rather than maximising likelihood, we maximise the log of likelihood. So, what we do, we will take log of this and then we will take the partial derivative of the same with respect to each parameter. So, if we have two parameters, we will have two equations if we have one parameter we will have an equation and these equations then, because two parameters two equations.

So, two equations when we solve you will get the value of the two parameters, but generally these equations tend to be nonlinear and once they are non-linear, in that case, it may result in a way to solve these many times, we may not be able to solve this by using the direct subtraction multiplication et cetera. We may have to use numerical methods to solve this sometimes we need the numerical methods and the most popular one or easier one is the Newton Raphson method. We can also use some other approaches like Quick Search et cetera and we are able to find a solution at which this value will be given the 0. We will not be in this this lecture series we are not going into much detail on exactly how to do this.

So, to solve these equations, generally, whatever statistical packages you are using mat lab Minitab or any other statistical packages, whichever has the facility of doing the maximum likelihood estimation they will themselves have the inbuilt functions that will give you the direct solutions. We will only solve wherever this solution is possible in a simpler way using the pen paper or simple excel sheets. So, here like if we take log of log likelihood this then what will happen if ln of L I am taking as an example.

So, what will happen this will become ln of pi of f ti let us say and pi of R ti where i is element of failure domain here i is for the censored domain also when we take inside ln what will happen this will become summation of i over failure that is submission of ln of f ti we know that multiplication of individual inside l log will be like ln of R1 into R2 is equal to ln R1 plus ln R2. So, same we have used here so, when we take inside ln that will become summation plus summation i as an element of R ln of R ti. So, similarly, when we have the particular function we can expand this further and we can find out the ln of L and then we can differentiate it. We will try to do this for one and let us so that you get an exposure that how this is solved.

(Refer Slide Time: 11:31)



So, let us see if we talk about the exponential MLE. So, for exponential MLE we know, we need to know what is f t, you know f t is lambda e to the power minus lambda t, that this we already know. And what is R t, R t is e to the power minus lambda t. These are the two functions which we have for exponential distribution. Now, for exponential distribution if you want to solve the same equation our likelihood is ln of l will become summation of ln of f ti where i is functional failure domain plus summation over i for censored data ln of R ti. So, this function we can use and we will be able to solve this, if I say this then this will be summation i equal i as an early where i belong to f, f ti is this.

So, that will become ln of lambda e to the power minus lambda ti plus summation I as an element of c ln of e to the power minus lambda ti. So, here if you see summation over i as an element of f. So, this will become ln of lambda minus summation i as an element of f lambda ti e to the ln of e to the power minus lambda ti will be lambda ti minus lambda ti same way this will become summation i as an element of c e to the power ln of e to the power will be minus lambda ti. So, as we see here this if I put it further then we get this equation that is ln of L. Now, with this equation whatever we see there is only one parameter lambda so, that let us say delta ln of L and we differentiate with respect to partial derivative across lambda.

So, this will become summation i for failure data that will be 1 upon lambda minus i for failure data this will be lambda into ti. So, this will become ti again, so minus is already taken here. So, this will be plus. So, here minus if you differentiate with respect lambda this will also become summation i n element of c ti this is the same equation which we were talking about. So, let us say if we have r number of failures for right censored let us say at time ts censoring happen of rest of the devices n minus r.

So, if that is the case, how much this value will be delta ln of L over delta this is used this can be used for multiply sensor data singly censored right censored every type of the data can be used here, delta lambda is equal to, now, how many see this i will be equal to i equal to 1 to r, because r failure is there. So, this will become r divided by lambda 1 upon lambda if I am summing up r times this will become r upon lambda minus summation of i equal to 1 to r ti that is if you see this is nothing but the summation of failure times minus summation i equal to, now here r failure have been there.

So, rest of the failure data which you have r plus 1 to n this data is having if you sum there that means are the censored times or I can say ti plus. So, and this has to be put equal to 0 for and if we solve this that means, r upon lambda is equal to summation i equal to 1 to r ti plus summation i equal to r plus 1 to n I am writing ti plus I am writing so, that you can differentiate that this is the data which is for censoring. So, this if we solve further we can know what is the lambda here lambda will be equal to r divided by summation of i equal to 1 to r ti plus summation i equal to r plus 1 to n ti plus.

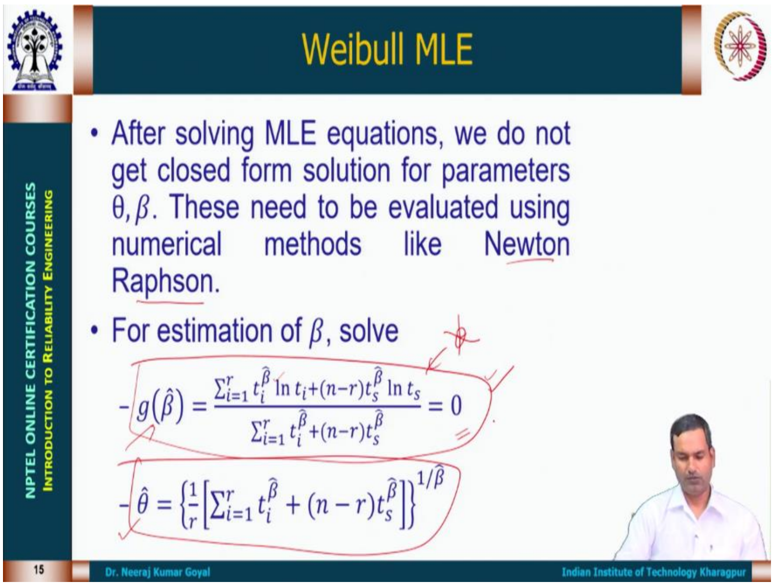
So, this is the formula which we have discussed lot about as we consider the observed failure rate as we calculated earlier. So, r is the number of failures and what is this, this is sum of failure times total failure times or cumulative failure times, and what is this data, this data is sum of operation time for censored data. So, as we discussed here, this total of failure time and total of operational time will become the total operation time, all together how much cumulative hours, or cumulative weeks, or cumulative days, or cumulative years, or device during test has built up, and this is the division factor for the number of failures, and this gives us the lambda.

So, lambda gives the r divided by t, where r is number of failures and t is the cumulative operation time which is submission of failure times and summation of the sensor times. Then if I say censoring is happening at n minus r then summation of ti plus for i equal to r plus 1 to n. Now, here this will n minus r is the total data points and all times ti plus or equal to ts that will become n minus r into ts. So, that will become the total time which is spent by the censored devices. So, whatever is the data you can use. This formula is actually because most of the time when we go into the practical applications, the inherently lot of calculations when we are doing we are assuming the constant failure rate or the exponential distribution.

So, when we are assuming exponential distribution, this becomes the formula to calculate the failure rate and MTTF is the 1 upon lambda. So, MTTF is calculated as MTTF estimated values 1 upon lambda cap that is equal to t divided by r that means total cumulative time divided by number of failures. But as we discussed if the assumption of exponential distribution is violated at the time this formula will not be valid. So, we should be careful when we are using these formulas. Similarly, I am not going in two more examples and solving these equations.

The reason being this we have already done like lambda, we can apply this formula and we can directly get the data we have the data from the data, we can and I have already solved one problem for this.

(Refer Slide Time: 20:18)



$$g(\hat{\beta}) = \frac{\sum_{i=1}^{r} t_i^{\hat{\beta}} \ln t_i + (n-r) t_s^{\hat{\beta}} \ln t_s}{\sum_{i=1}^{r} t_i^{\hat{\beta}} + (n-r) t_s^{\hat{\beta}}} = 0$$

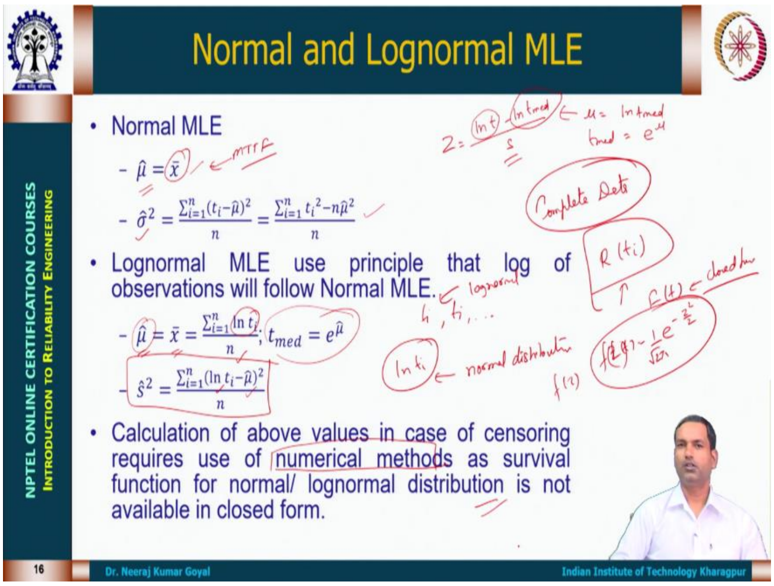
$$\hat{\theta} = \left\{ \frac{1}{r} \left[\sum_{i=1}^{r} t_i^{\hat{\beta}} + (n-r) t_s^{\hat{\beta}} \right] \right\}^{1/\hat{\beta}}$$

Similarly, for Weibull distribution, if you want to get the MLE that becomes little tricky, because for Weibull distribution, our two equations, which we will be getting, and when we solve, we are able to get fortunately here two equations, which are like equation one here, we are able to eliminate there is no theta here? This is without theta, same, once we do the differentiation with respect to beta, and theta, we will have the same process for t followed, we can follow here, I am not replicating that, because that will be a little bit more tedious looking and looking like a, which is not in we do not want to put it in here, but you can try it will not take much time, but you will be able to do it, you can verify whether you get the same value or not by following the same process was what you have followed earlier.

So here, if you see this, this is a nonlinear equation? Has it been linear equation, we could do the right side left side transformation, multiplication division subtraction, and we could get the value of beta directly. But since this is a nonlinear equation, we have to use numerical methods here like Newton Raphson. Or we can, we can sometimes also do the quick search where we try to have two values initially, and one should be giving the positive value for g B and other values should be giving the negative value. So that means the solution lies somewhere in between.

So, then we keep on changing the interval, or shorten the interval by tying the midpoint if midpoint is positive, then it replace the positive value side. If it is negative, then it replace the negative value side. So again, trying the midpoint, midpoint, and that is also we can reach to the solution or we can use the Newton Raphson here. Once you know the beta value here, then theta value can be directly calculated using this formula. But to calculation of beta, you have to use the Newton Raphson method or some other approach to get our guests values you have to put in try to do the hit and trial so that you are able to select a value beta which will give you this 0. This again, I am not going exactly how to do it we can do if you want but we are not covering it here.

(Refer Slide Time: 22:39)



Similarly, if we discuss about two other distributions like normal and lognormal for normal and lognormal MLE distributions also we can this is the straightforward formula which we generally use in the Excel sheet or anywhere like mu or the mean value is nothing but the actual mean value and sigma square that is the square of the second parameter is standard deviation that is nothing but the t minus mu whole square divided by n or we can say t square minus we can see here, MTTF mu is nothing but MTTF mean of time to failure if TTF is involved, so, this will become ti square minus n mu i square and mu squared divided by m. So, we can solve this accordingly. So, this gives us the sigma squared.

So, we can evaluate the mu and sigma directly using this formula. Similarly, like we know that for lognormal distribution, if you take the log of the time, it will follow the normal distribution. So, in a way, if we take the log of all the observations, and then we do the same exercise as we do for the normal, we will get the log normal distribution fitting that means, if all ti which I am getting here, t1 t2 if we take the ln of this, then ln of ti will follow the normal distribution, while ti is following the lognormal distribution. So, what change I have to do if I am writing this as a function like z is equal to ln of t minus ln of t median divided by s.

So, in that case, what happens I will whatever ti was using, so, rather than x bar I will take the ln of ti and average that means, all the ti data I will average or I will first do the logarithmic conversion and then I will use it. So, if I take average of that, that gives me the mu. So, what is mu here mu is ln of t medium mu is equal to ln of t median. So, how much is the median here the median here is e to the power mu. So, whatever mu you get if you take e to the power of mu you will get the t median and another parameter is the s is same as what we have here that is, but rather than ti again I will be using ln of ti ln of ti minus ln of t median which is nothing but the new cap. So, ln minus mu cap whole square by n the same formula when I use I will get the value of s.

So, these calculations they are able to do when complete data is there because for complete data same f t is applicable, but when data is censored, then as we have looked into that earlier for censored data, we need to know the R ti. This R ti is not known for the normal distribution, we know the PDF for the normal distribution, but we do not have the CDF for the normal distribution in a closed form. So, CDF is always evaluated in terms of numerical methods. So, if you want to use censored data here, then we have to again use numerical methods to solve the CDF value and from the CDF value to find out the value of theta, sorry, mu n sigma at which you will have the first derivative of the two functions as 0.

So, here because of lack of having the form f t is we do not have a closed form because we know that whatever we have f t or z t let us say we have z t e to the power minus z square by 2 1 upon under root to pi. Now, this function is non-integrable because it is non-integrable we are not able to have the CDF we only have the f z the only have the PDF the CDF is either determined by referring to the tables which have been numerically obtained or we have the function in software’s which use the numerical methods to give you this value. So, that is why in case of censoring it, we may find it difficult to use the MLE method for normal and lognormal, but it can be done the software's and if you use some programming language you can do that numerical methods and can do it you can do it in excel also.

But, in this class, we are not doing it maybe in some future class when we discuss this concept in more detail. We will try to do but here I do not want to go into that detail. So, because of this course has a limited it is an introductory course, so, we will not be going into much detail here. So, here like as we discussed today, we try to see some ways using which we are able to use MLE. But, as we have seen for MLE we have to solve the develop the likelihood equations, and then take the derivatives and then solve the equations which can sometimes turn out to be nonlinear equation.

But for complete data in case of normal lognormal we are able to have direct formula. Similarly, for complete or censored both type of data, we can solve exponential distribution and get the parameters of exponential distribution directly. But we may have to solve a nonlinear equation using the numerical methods for the Weibull distribution. We will continue have a discussion about it. And so we will stop here today. Thank you.