## **Introduction to Reliability Engineering Professor Neeraj Kumar Goyal Subir Chowdhury School of Quality and Reliability Indian Institute of Technology, Kharagpur Lecture 13 Normal Distribution**

Hello everyone. Now, we will start our lecture number 13 which is about Normal Distribution. So, in previous lecture we discussed about Weibull distribution which is a time dependent failure rate distribution. As we learned, when the shape parameter of the Weibull distribution becomes greater than 3, it tends to resemble a normal distribution. The normal distribution is an increasing failure rate distribution where the failure rate is also timedependent. So, today we will delve into the topic of normal distribution and discuss its characteristics in detail.

(Refer Slide Time: 01:04)



Normal distribution as we know it is very widely used and especially in quality. So, in quality almost all the formulas are driven by the inherent assumption that distribution is the normal distribution. So, normal distribution whoever goes through normal statistics estimation etcetera, they are very well versed that normal distribution is there.

So, inherent assumption is most of the time normal distribution. For a Lap T application a normal distribution can represent the stress distribution. It can represent strength distribution, it can represent wear distribution, it can represent the fatigue distribution and many other cases. So, normal distribution is used in reliability theory also, may not be as widely as weibull distribution for time to failure data.

But apart from time to failure, other parameters generally follow the normal distribution. The PDF of normal distribution is bell shaped whenever someone is asked to draw a distribution, they will default draw like this. Why? Because this is such a popular distribution, normal distribution, which is a symmetrical like whatever is the value at the same distance, we will have the same distance value on the left-hand side.

So, normal distribution is a symmetrical around the mean. This is the mean and this flex point is generally mu minus sigma. Similarly, this flex point is mu plus sigma. So, the distribution PDF Probability Density Function is given by small  $f(x)$  that is 1 upon under root 2 pi sigma e to the power minus half t minus mu divided by sigma whole square, where t is defined for all values from minus infinity to plus infinity.

$$
f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}, -\infty < t < \infty
$$

Generally, here, t is the random variable which can have both values, positive as well as negative. Mean is mean of, this distribution has two parameters mu and sigma. And here by the nature this mu is the mean and sigma is the standard division or sigma square becomes the variance. If we want to calculate CDF, for this that is f t, then f t is minus infinity to t f x d x.

Small f x is we have to replacing this. There is d x missing. So, this is d x. So, when we integrate this, we will get the CDF. This if I try to draw here, if this is mean, then at mean you will have value 0.5 and this will be something like this, you will see this more little later. Now, this function if you want like unlike exponential distribution, Weibull distribution which we could integrate to get the capital F t capital R t from small f x. But that is not feasible here, we are not able to integrate this.

$$
F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \left(\frac{t-\mu}{\sigma}\right)^2}, -\infty < t < \infty
$$

So, how do we get then F(t) value? To get the capital F t value there is a transformation. What we can do that we can transform the normal distribution into a standard normal distribution. What is the difference between the normal distribution and standard normal distribution?

Standard normal distribution does not have any parameter. This is our standard normal distribution, PDF. If you see here there is no parameter, it is directly function of Z. There is no other parameter which is not known here. But for normal distribution, there are two parameters mu and sigma which need to be known for normal distribution to be defined. How can we convert this into standard normal distribution?

(Refer Slide Time: 05:06)

Introduction Normal distribution is widely used in quality applications. For reliability applications, it is widely used for representing stress distribution<br>strength distribution, wear, fatigue etc. It's PDF is bell shaped curve which is symmetrical around mean. **NPTEL ONLINE CERTIFICATION COURSES**<br>INTRODUCTION TO RELIABILITY ENGINEERING  $f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2}, -\infty < t < \infty$ Mean is  $\mu$  and standard deviation is  $\sigma$ •  $F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \left(\frac{t-\mu}{\sigma}\right)^2}, -\infty < t < \infty$ This function doesn't have closed form solution. Therefore, this distribution is first verted into standard normal distribution by transformation  $z = \frac{t}{t}$  $\phi(z) = F(z)$  $2 \rightarrow$ - The PDF is  $\varphi(z) = \frac{1}{\sqrt{z}} e^{-\frac{1}{2}z^2}$  $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx = F(t) = 1 - R(t)$ The CDF is  $\alpha(z) = \int^z$ Failure Rate  $\lambda(t)$  =

To convert this into standard normal distribution, we consider this as Z. We consider that Z is equal to t minus mu divided by sigma square. Now, sorry not sigma square, this is sigma. So, when we put Z equal to this then differentiation of this d z over or d t will be equal to 1 upon sigma.

So, what happens that whatever this sigma which we had is removed because once we divide this by this 1 upon sigma this will be gone, as well as t minus mu open sigma becomes Z. So, this becomes e to the power minus half Z square into 1 upon under root 2 pi. This is because of the transformation loss.

$$
\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}
$$

$$
\varphi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = F(t) = 1 - R(t)
$$

So, we can remember that standard normal distribution phi Z is given as 1 upon under root 2 pi e to the power minus half Z square which is not having any parameter that this is why this is a standard. So, since this is a standard, we can take different values of Z and we can get the values of small phi Z.

But our concern is not to get a small phi z, our concern is to get the capital phi Z. So, capital phi Z is the integration of this small f x. Now, this is also not integrable directly. So, numerical methods are used and approximation methods are used which have been used to give the values of phi Z, capital phi Z which is the CDF.

But what happens here since this is a standard, it is not having any parameter, so for different values of z I can get the values of capital phi Z which is I can also say this is f z, CDF. Now, these tables if I show that there are standard tables available, so whenever we have the statistical tables, we can get these kind of tables, like this kind of table you will see.

(Refer Slide Time: 07:14)





You see this is a standard normal distribution where Z values are here and these are the CDF values. So, if I am interested that what is the value of CDF for minus 3 then minus 3.0 I will take and this my 0.00135, 0.00135 is the value which will become. If I am interested to know minus 3.02 then this will be my value 0.00126.

So, by having this standard table because there is no unknown, so different values of z, I can get the values of capital phi Z. I can also do this same thing by using the Excel. So, I will be using the Excel for this purpose in this lecture. But you can use both Excel also and if you do not have access to the Excel, you can use the standard normal tables.

Now, here if we see that we can get the CDF. How do we get? We will replace t minus mu divided by sigma as the Z and that will give us the value required for the PDF. And we know PDF is unreliability if this random variable is time to failure. So, reliability is equal to 1 minus f t or I can say this is equal to 1 minus phi Z.

$$
\lambda(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - \phi\left(\frac{t - \mu}{\sigma}\right)}
$$

Where Z is equal to t minus mu divided by sigma. Similarly, we know the failure rate formula. Failure rate is given as small f t upon R t. So, small f t is here and R t is 1 minus phi Z. So, we can use different formulas to get this failure rate for the normal distribution.

(Refer Slide Time: 09:11)



So, I have done this exercise in Excel. In Excel if you see, I have just put up all the values here for demonstration purpose. So, I have two parameters here mu and sigma. Now, let us see how does this change. Generally, if you look at the how this goes then if we look at it then if we look this small f x then t is minus mu. So, here mu is the location parameter.

Because if I change the mu what will happen f x will simply shift left or right. If I increase mu it will shift right if I decrease mu it will shift left. And sigma is the scale parameter here. Because if I increase the if let us say I make sigma 10 times then what will happen? I need to I will get the same value of f x if I take the t also 10 times and mu also 10 times. So, this has the impact of the changing the scale.

So, if I increase this then it will stretch and if I decrease then this will be this will be your contracted. So, that is why peakedness will be more if sigma is less and or it it will be having higher width if sigma is more. The same thing we can see from here. Let us see, let us first see how sigma change. Right now, I have taken sigma is equal to 0.2.

(Refer Slide Time: 10:49)





If I take sigma is equal to let us say 0.5. Now, let us see what will happen to small f x. If you see small f x has become wider. Then I take this as 0.1, it will become narrower, sharp. So, generally, we desire sigma to be lesser for any distribution. Why? Because lesser the variability or lesser will be the uncertainty and it will be easier for us to make the decision.

More uncertainty means there is a more variability and because of more variability for it is difficult to decide that what is the right value and we have to consider the full range of the possibilities. So, here, same way, if I let us say, if I shift, if I make mu then it should be able to shift it.

## (Refer Slide Time: 11:44)



So, if I make it let us say from 1 I will make it 1.2. If I make it 1.2, see this mean is just shifted. If you see if I make it 0.8 then it will be shifted left. So, mu acts as a location parameter and sigma acts as a scale parameter. And if I increase the value of sigma, the width of the distribution will be more.

# (Refer Slide Time: 12:19)





If you see it here this is the Z X function which is increasing like if the whatever combinations I have taken, it is always increasing. If I take 0.1 And if I take this as let us say 2. So, shows the either similar or increasing. It generally does not decrease. So, here as we can see that normal distribution is symmetrical distribution around the mean and this has these properties.

Let us take an example. For normal distribution lot of literature is available or more you want to know about it, you can study more. Here we are not discussing much in detail, we are just looking at it that how does it, how we can use it for the calculation purposes.

#### (Refer Slide Time: 13:30)



So, let us see if we have a, let us take one example that there is a wear out. Wear out out means something which is let us say tire. So, tire what happens over the time they will keep on reducing the material it will keep on wearing out. So, here oil drilling bit is there. So, like our tools, mechanical tools which we use for drilling etcetera over the time they will lose their material.

So, that will continuously wearing out as we use. So, let us say this wear out is normally distributed with mean of 120 drilling hours and standard deviation is 14. So, what we can say here that our 120 drilling hours is the mean. So, mu is equal to 120. And we have standard division that is sigma, sigma equal to 14 hours. Now drilling occurs 12 hours each day.

So, in one day only 12 hours it works. How many days should drilling continue before operation is stopped in order to replace the drill bit. A 95 percent reliability is desired. What does it mean? We have to decide a interval here that if I start using the bit from time t equal to 0 then how long should I work so that I am getting the 95 percent reliability because I am taking 5 percent, 5 percent chance here that the drill may fail but 95 percent of the chance the drill will not wear out during this period.

$$
Pr\left\{z \ge \frac{t_{0.95} - 120}{14}\right\} = 1 - \Phi\left(\frac{t_{0.95} - 120}{14}\right) = 0.95
$$

So, I want to know the life for which reliability is 0.95. So, I see here that reliability is equal to 0.95 and that is equal to 1 minus phi t minus mu divided by sigma. Now, I want to know what is the value of t for this. So, I can solve this repeated in a reverse manner. So, phi t is 0.95 minus mu divided by sigma is equal to 1 minus 0.95. That means my failure probability is 0.05.

Now, here we want to know that how, this is in hours, mu is in hours, sigma is in also operating hours. So, whatever result we will get that of 40 this will also be in hours. So, I can do this t 0.095 minus mu divided by sigma is equal to phi inverse. Phi inverse is the inverse distribution, inverse value of the normal distribution for 0.05. This I can get in multiple ways.

This value I will show you how we can get it. Let us say for now that this value turns out to be minus 1.645. How do we get this value? I will show you. So, t is 0.95 minus mu divided by sigma is this value. So, t 0.95 would be equal to, mu will go here that will be mu minus 1.645 into sigma.

$$
\frac{(t_{0.95} - 120)}{14} = -1.645, \text{ or } t_{0.95} = 96.97 \text{hr}
$$

So, this value once we supply here then this value comes out to be 96.97 hours. Now, this is in hours. I want to convert it into the days. So, I have to divide this by 12 and that will give me the days. So, that comes out to be approximately 8 days. Now, let us do this calculation using either, first let us see how can we do this using the Excel sheet.

(Refer Slide Time: 17:42)











So, I have put an Excel sheet here. So, let us say mu, mu value is here 120. We have sigma or S D, standard deviation, standard division is 14. As we discussed earlier here we want to know the phi inverse for 0.95, 0.05. So, Z value will be equal to norm inverse, phi inverse means norm inverse, norm inverse. We will use norm inverse, norm inverse of what probability 0.05. And what would be the mean?

This is mean. Generally, this gives directly. If we want to solve with standard normal distribution then we will first take the zero mean, standard normal distribution has 0 mean and sigma s unity. If you replace mu as 0 and sigma as 1, the normal distribution will become standard normal distribution. So, let us do this. Now, if you see that we got the same value minus 1.645.

Now, since we have got this Z. Now, I want to calculate the t. t will be equal to as we discussed this will be equal to Z multiplied by sigma plus mu that turns out to be 96.97 hours. I want to get this in days. So, that will be equal to this divided by 12. That is approximately 8 days.

$$
\Pr\left\{-1.645 \le z \le 1.645\right\} = 0.90
$$
\n
$$
\frac{25,000 - \mu}{\sigma} = -1.645 \frac{35,000 - \mu}{\sigma} = 1.645
$$
\n
$$
\mu = 30,000 \sigma = 3039.5
$$
\n
$$
R(24,000) = 1 - \Phi\left(\frac{24,00 - 30,000}{3039.5}\right)
$$
\n
$$
= 1 - \Phi(-1.97) = 0.9756
$$

We have done this in two step by first getting into the standard normal than normal. We could have also got it directly. How we could have got directly? We could have got it by norm inverse. Norm inverse of 0.05 with mu this and sigma this and this will give us the value directly 96.972 hours. This is how we can solve this problem using the Excel.

(Refer Slide Time: 20:22)











Now, if we want to solve this use, as we see, let us see, if we can use the PDF table also. If we look at the table, our interest is to get the Z inverse for 0.05. Phi value is 0.05. Let us see what is the phi value here for 0.05. So, let us search here, which value is coming close to 0.05. This is 0.01, this is 0.48, this is 0.46. If we see here, sorry 0.05 we required. 0.06, 0.05 is coming somewhere here. 474849, if you see this. So, what is this value here?

This value is coming as minus 1.6 and here it is 4. So, either I can say minus 1.64 or I can say minus 1.6, somewhere in between minus 1.64 and minus 1.65. So, I can say minus 1.645 which is same minus 1.645. This value, this table gives us only the Z value. So, we get this Z value then we can get the value of t by multiplying this Z value, Z value with sigma and adding the mu value to this that will give us this 96.97 hours. Let us continue with our discussion with next example.

(Refer Slide Time: 22:06)



Let us say we have another problem where we say that 5 percent of a certain grade of tires wears out before 25,000 miles. That means within 25 miles we have a wear out of 5 percent and another 5 percent of the tires exceeded 35,000 miles. So, 5 percent is below 25,000 and 5 percent is above 35,000.

Now, we want to know the tire reliability as 24,000 miles, if wear out is normally distributed. Now, this problem if you look at it then let us first see in this presentation form that what does it says, it says that what is the probability if we draw the f x. Then we say that for 25,000 miles the chances are 5 percent. The area under the f x curve is 5 percent.

Similarly, for above 35,000 miles, the area is 5 percent. So, since we know that this distribution is having the symmetrical property around the mean. So, since 5 percent this side is 25,000 and 5 percent of 35,000 is this side. So, how much will be the mean here? The mean would be we can directly get it the mean would be mu would be equal to 25,000 that is the middle point of this.

Because if you see here, this is exactly related same distance from here to here. So, that is the middle point of this 25,000 plus 35,000 divided by 2. So, we get the 30,000. But we can solve it in another ways. Let us say Z 1 corresponding to this will be equal to 25,000 minus mu divided by sigma and this is our Z 2. Z 2 is equal to 25,000, sorry, 35,000 minus mu divided by sigma.

And we know that probability that Z is less than equal to  $Z$  1 is equal to 0.05 and probability that Z is greater than Z 2 is equal to 0.05. This we can also write as 1 minus probability that capital Z is less than Z 2, that is equal to 0.05. So, here by using this now or we can look at the table directly.

Now, so, this we can see probability that  $Z$  is less than equal to  $Z$  2 will be equal to 1 minus 0.05 that is 0.95. Now, if we look at the table, we have already seen that for 0.05, the value was minus 1.645. Similarly, if you look at the table again for value 0.95 because of the symmetry, this is plus 1.645. So, we know. Now,  $Z$  1 is equal to this and  $Z$  2 is equal to this.

So, here we can say that 25,000 minus mu divided by sigma is minus 1.645 and thirty 35,000 minus mu divided by sigma is plus 1.645. Now, if we solve this, we can get the value easily like that is we can write it as 25,000 minus mu is equal to minus 1.645 sigma. And another equation is 35,000 minus mu is equal to 1.645 sigma. If I subtract this first equation from here, so then what will happen?

This will become minus, this will become plus, this will become plus. Now, what will happen? Or I can subtract directly here also. So, if I subtract this. So, that will become 10,000 mu minus mu will be equal to 0 and this minus this will be 1.645 into 2 into sigma. So, I can get sigma as equal to 10,000 divided by 1.645 into 2. And similarly, if I want to get the mu I will sum it up.

So, that will become 60,000 minus 2 mu will be equal to 0. So, two mu is equal to 60,000 and mu is equal to 60,000 divided by 2 that is 30,000. So, I am able to get mu will also I am get able to get the sigma also. Once I get this, my parameters are now known. Since I know the parameters mu and sigma, I can calculate any probability I want.

The probability which I want to know is the reliability for 24,000 hours. So, let us first calculate reliability for 24,000 hours. Reliability is 1 minus CDF 1 minus CDF, phi of 24,000 hours. So, this is 24,000. So, 24,000 minus mu divided by sigma, 24,000 minus 30,000 divided by sigma 3039.5. Now, this if I look into the table, this value comes out to be minus 1.97.

(Refer Slide Time: 28:25)



It is minus 1.97 value in that table if you look at it minus 1.97. So, I will look into this minus 1.9 and 1 0 1 2 3 4 5 6 7, 0.02442. So, 0.02442 comes out to be the value of phi. And if I subtract this value from 1, the value which I will get is 0.9756 approximately.

## (Refer Slide Time: 29:00)



So, we can use table or we can use this also directly to get these values, that is equal to my mu which I have calculated is 30,000 and standard deviation which I have got is 3039.5. So, I can get the R 24,000. R 24,000 is equal to norm 1 minus 1 minus norm distribution. I will use norm distribution here.

Norm distribution for x, x is 24,000, 24,000 and mean is 30,000 and sigma is this. And I want the cumulative value. So, I will use the true here. If I do this I get the probability 0.97581 directly. I do not have to use the first standard normal and then convert into the normal. I can directly also calculate using this Excel sheet.

## (Refer Slide Time: 30:31)



There is an important theorem which is used with respect to normal distribution that is if we use different distribution, different random variables and if you take summation of those random variables as y then if these distributions they may be different distribution, they may not follow the normal distribution.

If they follow normal distribution then for any number small n, whatever is the number, 5, 10, this will always be the normal distribution. But central limit theorem states that irrespective of the distributions of  $x1 \times 2 \times n$ , if n is very very high, n is approaching infinity then summation of these random variables, whatever random variable we get, this random variable will follow the approximate normal distribution.

And the mean value, we have the two parameters for normal distribution, mean and standard deviation. So, mean of this random variable is nothing but the summation of the mean of the random variables x1 to xn and mean of this output random variable, variance of this output random variable is nothing but the summation of variance of the individual random variable.

So, like if we use a standard, if we use x, if we let us say these are following the exponential distribution then mean is 1 upon lambda. So, if all are same, let us say then this will become n upon lambda. Mu will be equal to n upon lambda and variance will be equal to, because this is 1 upon lambda square that will become n upon lambda square.

Any other distribution whatever is their mean, whatever is their variance, we put it here to calculate the mean and variance of the resultant random variable. Now, since we know the mean and variance of the resultant random variable, we know the distribution of random variable and we can use it to get any quantity of the interest. We will stop here today and we will continue our discussion with the next distribution would be log normal distribution. Thank you.