

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture – 13
Continuous Probability Distributions (Part 2)

Once again hello all of you, so in continuation to our earlier discussion on continuous probability distributions where we have studied uniform probability distributions, normal distribution and standard normal distribution. Now we will be studying some more continuous probability distributions.

(Refer Slide Time: 00:42)



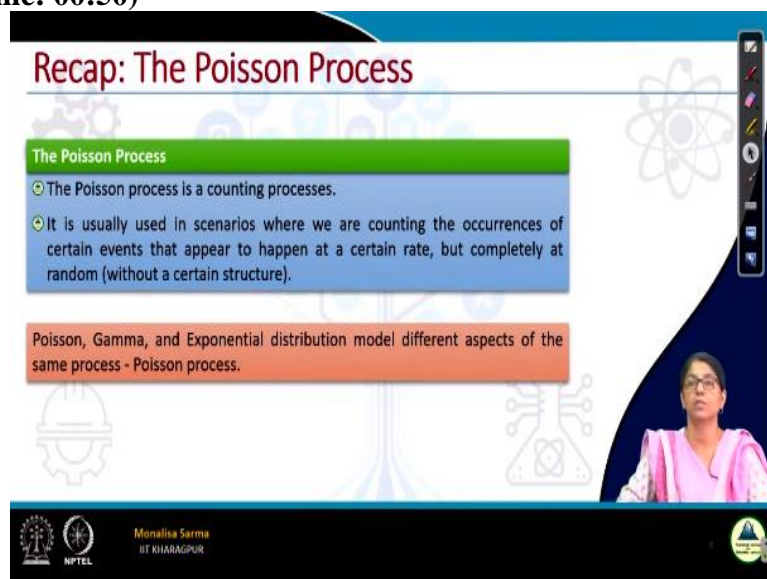
The slide is titled "Concepts Covered" and lists three items:

- Gamma distribution
- Exponential distribution
- Lognormal distribution

The slide also features a video feed of Prof. Monalisa Sarma in the bottom right corner and logos for NPTEL and IIT Kharagpur in the bottom left corner.

So, in this lecture basically we will be studying gamma distribution, exponential distribution and lognormal distribution.

(Refer Slide Time: 00:50)



The slide is titled "Recap: The Poisson Process" and contains the following text:

The Poisson Process

- The Poisson process is a counting processes.
- It is usually used in scenarios where we are counting the occurrences of certain events that appear to happen at a certain rate, but completely at random (without a certain structure).

Poisson, Gamma, and Exponential distribution model different aspects of the same process - Poisson process.

The slide also features a video feed of Prof. Monalisa Sarma in the bottom right corner and logos for NPTEL and IIT Kharagpur in the bottom left corner.

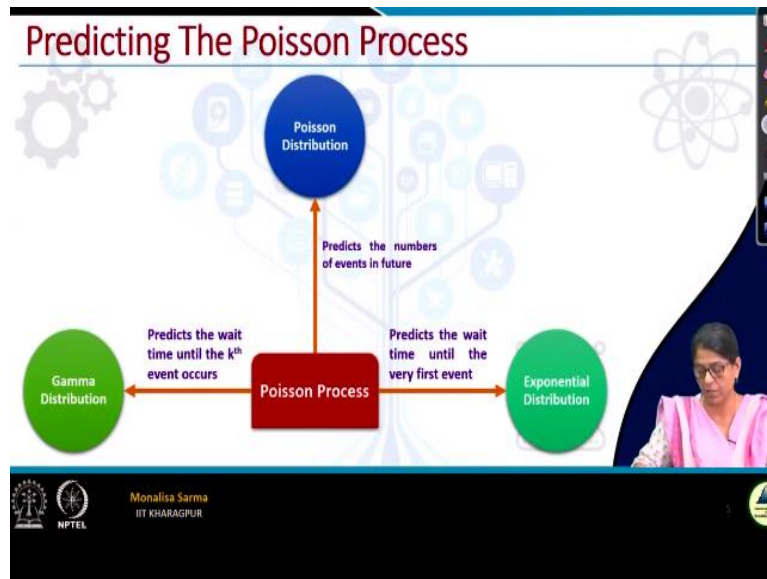
Now, before going to gamma distribution, let us take a quick recap on the Poisson process. Now, why am I suddenly bringing the Poisson process you will see? So, what was the Poisson process if you can remember Poisson processes was nothing but it was a counting process. Basically, I can say it is usually used in scenarios where we are counting the occurrence of certain events that appear to happen at a certain rate where the average occurrence rate is given.

So, when average occurrence rate is given, we are interested in the occurrence of a certain number of events within a time interval or within a region of space is not it? But this occurrence is completely at random. Like if let me take you an example suppose we know the occurrence of earthquake in a particular area, so particular area say maybe in square meters or whatever it is in occurrence of earthquake is given at a particular rate, say around in one year say around 5 earthquake has happened. So that is the average rate.

Now, we know in average occurrence rate is 5% 5 earthquake, but does not mean that we will know exactly this when earthquake will happen within an interval within in 1 year. Moreover, if in 1 year 6 earthquake happened, that does not mean that in the next interval, since the indices or 6 earthquake has happened that does not mean in the next interval next year, maybe we will get less number of earthquake it is not that as well.

So that we have seen that was the property of memory-lessness so that is what this occurrence is totally random, that was Poisson process we have seen. Now, why I brought out this poisson distribution here because Poisson distribution, gamma distribution and exponential distribution, they model different aspects of the same process that is a Poisson process, that is the counting process, they model the same aspect but in a different way.

(Refer Slide Time: 02:46)



How and Poisson distribution as we have seen, we predict the number of events in the future, given the average occurrence rate, we know what is the how many events is on an average the number of events that occur within a time interval, then based on that we can predict a number of events that may happen within it in a time interval. That is what we do in Poisson distribution. Whereas exponential distribution we say exponential distribution, we try to predict the occurrence of an event.

Maybe the occurrence of a random event, maybe the occurrence of a failure when you talk in reliability studies, I maybe talk a failure occurrence a failure of a component or maybe the random event maybe like my bus coming in coming / bus at a particular time. So, basically it predicts the occurrence of an event or rather I should say it predicts the wait time until the very first event.

So, when is the time that the first event has occurred or we can also the data say time between 2 events. These, all these are can be modelled by exponential distribution. Now, what is the gamma distribution? Gamma distribution predicts a wait time until the k th event occurs. So, see the difference Poisson distributions given the average rate, we can predict a number of events that can happen in the future exponential distributions, we model the wait time between for the first event or the wait time between 2 events.

Or we can also say occurrence of an event it models the occurrence of any random event, this gamma distribution, it predict the wait time until the k event occurs. The time that pass so that k event occurs. Now first we will be discussing gamma distribution.

(Refer Slide Time: 04:43)

Gamma Function

The Gamma distribution derives its name from the well known **Gamma function** in mathematics.

Gamma function

Gamma function is given by,

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{for } \alpha > 0$$

Integrating by parts, we can write,

$$\begin{aligned} \Gamma(\alpha) &= (\alpha - 1) \int_0^{\infty} x^{\alpha-2} e^{-x} dx \\ &= (\alpha - 1)\Gamma(\alpha - 1) \end{aligned}$$

Thus Γ function is defined as a recursive function.

Monalisa Sarma
IIT KHARAGPUR

Gamma distribution, this name it has got from the gamma function, gamma function most of you I think you already know what is a gamma function. It is basically what we have factorial for integer numbers, this gamma function it is a representation of that only, but it is in the complex numbers and real number for factorial is always for integer number. So, gamma function is a representation of that, but in complex and real numbers.

So, gamma function we can specify this is the gamma function. So, this is integration means, if we will have to do integration by parts I told you on one of the class will have to learn integration so, how to do integration the part so, if integration the part so, we get gamma of alpha is this is the thing that means, if you are interested in finding Γ of 3 that means what is Γ of 3 that is 2 into Γ of 2 and what is Γ 2? Γ 2 is 2 into Γ 1.

(Refer Slide Time: 05:49)

Gamma Function

Properties of Gamma Function

- When $\alpha = n$, we can write,
$$\begin{aligned} \Gamma(n) &= (n-1)(n-2) \dots \dots \dots \Gamma(1) \\ &= (n-1)(n-2) \dots \dots \dots 3.2.1 \\ &= (n-1)! \end{aligned}$$
- Further, $\Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Monalisa Sarma
IIT KHARAGPUR

So, I have also here some of the properties of the gamma function. So, if $\alpha = n$ then n means, if it is an integer then we can write it is $n - 1$ Γ of n is $n - 1$ factorial whereas, factorial of n is n factorial. So, Γ of n is $n - 1$!, for the gamma of n is nothing if you can put there instead of n if we put the value we will get in the expression the last expression what we got here if we put the value $\alpha = 1$.

So, we will get Γ of 1 is 1, $\Gamma(1/2)$ is $\sqrt{\pi}$ these are some standard things which you can remember and of course, you can find out a gamma function for any given value. So, from this my gamma distribution has this function in it the gamma density function that density function corresponding to gamma function has this function in it and so, we call it as a gamma distribution function.

(Refer Slide Time: 06:44)

The slide is titled "Gamma Distribution" and contains the following information:

- Formula: Gamma Distribution**
- The continuous random variable x has a gamma distribution with parameters α and β such that:
- The probability density function is given by:

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & \text{Otherwise} \end{cases}$$
 where $\alpha > 0$ and $\beta > 0$.
- A graph shows the probability density function $f(x)$ for three different values of α (with $\beta = 1$):
 - $\alpha = 1, \beta = 1$: A single peak at $x = 1$.
 - $\alpha = 2, \beta = 1$: A peak at $x = 2$.
 - $\alpha = 4, \beta = 1$: A peak at $x = 4$.
- A handwritten note in red says $\alpha = 1$.

So, gamma distribution function has 2 parameters again here what are the 2 parameters? α and β . So, α is basically the shape parameter like till now, we did not get the shape parameter normal distribution did not have the shape parameter here we have to separate the way you see the see for different values of α we get differentiate. So, here for $\alpha = 1$ this shape $\alpha = 2$ we know got this sort of shape this one. This $\alpha = 2$ and $\alpha = 4$ we got this shape.

So, we got different shape for different values of α . So, α is that is why it is called a shape parameter. So, now, what is β here? β is the rate of event occurrence. So, if the β is rate of even occurrence that means, what is in terms of λ , what is λ in this Poisson distribution we have λ . λ is a number of events that occur within a time average number of events that occur

within a time interval. So, if the average number of events that occur within a time interval is α , suppose x events occur within a time interval.

So, what is the thing rate of the occurrence of event? Rate of the occurrence of event will be α . β is nothing but $\beta = 1 / \alpha$. So, here like exponential distribution so, this is the density function for exponential distribution basically, now here are sorry this is the density function for gamma distribution. Now, exponential distribution is a special case of gamma distribution special case in the sense like in this gamma distribution.

If we put $\alpha = 1$ resulting distribution is exponential distribution. So, here in this figure this $\alpha = 1$ this is exponential distribution this is the curve corresponding to exponential distribution. So, this is the exponential distribution we call it this special case of gamma distribution. So, exponential distribution again it is, there is just 1 shape so, there is no shape parameter value of α is just 1.

(Refer Slide Time: 09:05)

Exponential Distribution

Definition: Exponential Distribution

The continuous random variable x has an exponential distribution with parameter β , where:

$$f(x; \beta) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & \text{where } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Note:

- The mean and variance of gamma distribution are

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$
- The mean and variance of exponential distribution are

$$\mu = \beta$$

$$\sigma^2 = \beta^2$$

Monalisa Sarma
IIT KHARAGPUR

So, exponential distribution what is the formula same in the gamma distribution probably density function putting $\alpha = 1$ whatever we get that is definitely see here if I put $\alpha = 1$ what do I get $\alpha = 1$ $\Gamma(\alpha)$ $\Gamma(1)$ is 1, so, β^α that is β then $x \alpha^{-1}$ that is $1 - 1/x$ tends to the 0, so, $e^{-x/\beta}$ so, what I got $1 / \beta e^{-x/\beta}$ is not it? Here if I put $\alpha = 1$ in this expression if I put $\alpha = 1$ what do I get?

I get $1 / \beta e^{-x/\beta}$ is not it? That is the probability density function for exponential distribution. See, that is the probability density function for exponential distribution now what is β ? β is

nothing but it is the rate of occurrence or event rate of occurrence of event means, if the average number of events that occurs within a time and time interval is λ so that means β is $1 / \lambda$.

So, this if I write it in terms of λ that is the parameter of Poisson distribution λ was the parameter Poisson distribution is not it? Λ gives the average number of events that happens within a time interval. So, if I write this expression in terms of λ , so, what it will be my expression becomes $\lambda e^{-\lambda x}$ if I write in terms of λ . So, now the mean and variance of gamma distribution is mean is $\alpha \beta$ variance is $\alpha \beta^2$.

Now, if this is the mean and variance of gamma distribution, so what will be the mean and variance of exponential distribution then, from this exponential we will be able to say how putting $\alpha = 1$. So, if we put $\alpha = 1$ we get $\mu = \beta$, σ^2 as well a β^2 that is all. So, mean and variance of exponential distribution is β , β or we can tell also $1 / \alpha$ and variance is β^2 that means, standard deviation σ is β .

So, mean and this is also one important characteristic which is to remember mean and standard deviation of exponential distribution is seen that is β or $1 / \lambda$ mean and standard deviation and variance mean and standard deviation of exponential distribution is $1 / \lambda$ or we can say β .

(Refer Slide Time: 11:42)

Reliability Function of Exponential Distribution

Derivation of the Reliability Function of ED

The pdf of Exponential distribution:

$$f(x; \beta) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & \text{where } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Here, β = mean time to occurrence of an event

We can write $\lambda = 1/\beta$, where, λ = mean number of events per unit time

From the pdf of exponential distribution, we get $f(x; \lambda) = \lambda e^{-\lambda x}$

The cumulative distribution function is given by $F(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}$

Handwritten notes: $F(x) = 1 - e^{-x/\beta}$
 $R(x) = 1 - F(x)$

So, see here I have mentioned where λ is mean number of events per unit time. So, we can write λ is equal to where β means mean time to occurrence of an event, this is mean number

of events per unit time, say remember the difference. Now, this is $F(x)$ this is the probability density function. So, if you are interested in finding the cumulative distribution function like how we have found it for normal.

Normal from what to say integrating from $-\infty$ to the particular value we are interested similarly, we can find out the cumulative distribution function for gamma also gamma distribution also exponential, exponential distribution also now, let us see for exponential distribution, if you are interested in finding out the cumulative distribution function. So, what it will be $F(x)$ is integration of 0 to $x \lambda e^{-x\lambda}$ So, what we got this is the value.

Now, one thing this exponential distribution normal distribution, all this distribution has gamma distribution, this is where it used in reliability studies. So, when we use in reliability studies, we basically do not use x as a random variable. The general convention basically instead of x we write t because in reliability and reliability studies, we basically try to model the, our random variable is the time, time to failure time to repair whatever it is. So, when we talk of time, so in the variable, instead of using the variable x , we use the variable t that is all.

So, that is just a general convention that is all. So, now, this is the cumulative distribution function $1 - e^{-x\lambda}$ just simple integrating the probability density function, so, if this is $F(x)$, now, let us what is reliability? Reliability is the probability of failure free operation is not it? So, what is $F(x)$? $F(x)$ is the probability that it will fail from $-\infty$ to x probability that it will fail till x th time what is x ? x is nothing but if I talk about reliability it will be in terms of t in terms of time.

So, what is $F(x)$ instead of writing $F(x)$, let me write this $F(t)$, what is $F(t)$? $F(t)$ the integration from 0 to t means cumulative failure probability of failure till time t . So, probability of failure till time t I got $1 - e^{-\lambda t}$, this is my probability of failure. Now, if this is my probability failure then what is the reliability? Reliability is probability of failure free operation this is F_t normal reliability is nothing but 1 minus of F_t .

Similarly, this here I have shown in exponential distribution same case you can apply for all normal distribution. Normal distribution also you can find out F_t then R_t is 1 minus of F_t . F_t how same integration from 0 to t using the probability density function that is F_t then $R(t)$ is 1 minus F_t this of course, I mentioned it here, but you can use the same process what will be

the difference and nothing different. So, it is the same process we can apply for all the other distribution as well. So, my reliability is $e^{-x\lambda}$.

(Refer Slide Time: 15:35)

Exponential Distribution Properties

- 1 The exponential distribution has the memoryless property.
- 2 Memory-lessness occurs when failure is caused by external causes only
- 3 This implies the failure of a device due to random or sudden shocks, but not due to deterioration or wear.
- 4 So, it imparts an item a property of being as good as new.
- 5 Failures of electronic and electrical components obey this distribution.

Monalisa Sarma
IIT KHARAGPUR

Now, this exponential distribution it has a property one of the most important property is that it has a memory-lessness property what is this memory-lessness property in case of exponential distribution like, what does exponential distribution failure? Exponential distribution it has a constant failure meaning like suppose if I talk about vehicle, vehicle is not a very in the normal useful remember when I talked about the failure rate that is the fact of that.

First usually it is the quality failure we say that is the infant failure rate then is the normal working life as a normal failure then is the wearout failure is not it? In a normal working life the failure that happens is maybe due to random reasons, this may not be because of some problems in the machine may not be some quality related failure it may not be because of wearout failures in a normal working life when a failure happens it failed due to some random reasons.

So, in this time the failure that happens is because of it has a constant failure rate. So, constant failure rate meaning is that so, suppose if I talk of a vehicle only in its useful life that suppose a vehicle that has failed at time $t = 100$. So, now at $t = 100$ a vehicle that has failed, it will not be different if I consider that the vehicle has run for 1000 kilometre before that, or the vehicle has run 1000 kilometre before that just because the vehicle has run 1000 kilometre before that does will have any affect that.

If I tell that I am at a certain point of time, if I consider what will be the probability that it will fail after 100 hours. Suppose the vehicle has run 1000 kilometre and I am asking so what is the probability that it will fail after 100 hours? Again, there is another vehicle which has run say 100 kilometres only then I am asking what is the probability that it will fail after 100 kilometres if this failure was a failure can be represented by exponential distribution.

So, if this failure shows a constant failure rate, then whether it run 100 kilometre or a it run 1000 kilometre it makes no difference probability that it will fail after 100 hours in both the case will be same this happens why this sort of failure will only be seen in case of random events so random events may not unless it is written here. So, memory-lessness this failure x what is it exponential distribution like this constant failure rate?

It happens when failure is caused by external causes only maybe due to some external causes, but maybe the external causes. So, this implies that failure of a device due to random or sudden shocks but not due to deterioration or wear maybe due to sudden shock. Then so, it imparts an item the property of being as good as new whether it has 1000 kilometre or it has an 2000 kilometre does not make any difference.

Failure of electronics and electrical components obey this type of distribution, this exponential distribution it always electronic and electrical components obey this distribution even for software failure also software failure we see this type of distribution when we talk about software here or only we are talking of hardware of course software also we can think of software reliability. So, when we took me think of software failure.

Its software failure also what to say it can be model using exponential distribution here also the failure rate is due to some random event here random for a software random event not been may not be due to environment maybe some other causes, which I will not dwell into details.

(Refer Slide Time: 19:27)

Exponential Distribution—Example 1

Problem

Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse by the time 2 calls have come in to the switchboard?

$\lambda = 5$
 $\beta = \frac{1}{5}$

So, we will see some examples here. Suppose that and telephone calls arriving at a particular switchboard follow a Poisson process see here, it follows a Poisson process. Poisson process means it is a counting process, basically. How many calls arrived at a particular switchboard? That is why is calling the Poisson process. So, we are interested in the events, different events. So, what is with an average of 5 calls coming per minute, so per minute, it is 5 calls that means my λ is 5 calls.

It my λ is 5 then what is my β ? β will be equals to $1 / 5$. So, here it is given 5 calls coming per minutes that means my $\lambda = 5$ and Poisson distribution we have seen this is not it? So, now what is my β ? β will be $1 / 5$. So, what is the probability that up to a minute will elapse by the time 2 calls have come into the switchboard? See here what it is asking, what is the random variable here?

Random variable is definitely the time always I am repeating again which I repeated in my last few classes also always try to find out what is the random variable. So, what is the random variable here? The random variable is the time but it is asking time up to a minute with the last by the time we need, what is the probability that 2 calls have come? So, my number of calls is 2 my number of call is 2 and it is telling up to a minute what is the probability that 2 calls will come?

If number of calls is 2 that means see if you see the first slide well I have to explain the difference between the Poisson process let me come to those that slide this one see here. Poisson exponential and gamma, exponential is the first event order time between 2 events.

What does gamma says predict the wait time until the kth event occurs? So, wait time and kth event so here kth event from 0 to 2 events. So that means this question is a question of gamma distribution.

(Refer Slide Time: 21:37)

Gamma Distribution-Example 1

Problem

Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse by the time 2 calls have come in to the switchboard?

Solution

The Poisson process applies, with time until 2 Poisson events following a gamma distribution with $\beta = 1/5$ and $\alpha = 2$. Denote by X the time in minutes that transpires before 2 calls come. The required probability is given by

$$P(X \leq 1) = \int_0^1 \frac{1}{\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx = 25 \int_0^1 x e^{-5x} dx = 1 - e^{-5}(1 + 5) = 0.96$$

Monalisa Sarma
IIT KHARAGPUR

So, it is a question of gamma distribution. So, simple whatever is given it is now it is a simple question we know that it is equation on gamma distribution, we know the value of β what is the value of β ? B value is $1 / 5$ we know what is the value of α ? $\alpha = 2$ just simple put it in the expression; put it in what to say expression for gamma distribution?

(Refer Slide Time: 22:21)

$f(x, \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$

$\alpha = 2$

$\beta = 1/5$

$x > 0$

$\frac{1}{\beta^2 \Gamma(2)} x^0 e^{-x/\beta}$

So, you remember what was the expression for gamma distribution it is $1 / \beta^\alpha \Gamma(\alpha) x^{\alpha-1}$ then it is $e^{-x/\beta}$ where $x > 0$ this is $f(x, \alpha, \beta)$ this is the formula is not it? Now, here α is given 2. So, what is $\Gamma(2)$? So, α is given 2 β is we got β is equals to $1 / \alpha$, α is 5 so β is what? $\beta = 1 / \alpha$. So, α is 5 so β is $1 / 5$ is not it?

So, afterwards we will take substitute a value for β or first let us have to value of 2α and that is α is 2 and $\Gamma(2)$ what will be to Γ of 2? $\Gamma(2)$ will be 1 $\Gamma(2)$ will give us 2 into 1!. So, it is basically 1 into 1!, so $\Gamma(2)$ will be 1 so x of $\alpha - 1$ so $2 - 1$ that is x of 0, x of 0 is 1. So, $e^{-x/\beta}$ so, this is the expression. So, here this is the expression.

So, simply by solving it whatever we got that is the probability that up to a minute will elapse by the time which we will get 2 events. See the difference between the Poisson distribution and gamma distribution. In Poisson distribution what we tried to find out what was the probability that 3 events have occurred, what is the probability that 4 events have occurs in a particular time? There are random, what is the random number? Random number was the number of events.

Time was fixed and we wanted to find out the number of events within that time interval. Now, here what is the random number here? Random number is time t , in time t we are interested in finding out what is the probability that in time t 2 events is occur. The random number is the time and in the interval.

(Refer Slide Time: 25:02)

Exponential Distribution—Example 2

Problem

Suppose that a system contains a certain type of component whose time, in years, to failure is given by T . The random variable T is modeled nicely by the exponential distribution with mean time to failure $\beta = 5$. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

Solution

The probability that a given component is still functioning after 8 years, is given by

$$P(T > 8) = \frac{1}{5} \int_8^{\infty} e^{-\frac{t}{5}} dt = e^{-\frac{8}{5}} = 0.2$$

Let X represent the number of components functioning after 8 years. Then using the binomial distribution, we have

$$P(X \geq 2) = \sum_{x=2}^5 b(x; 5, 0.2) = 1 - \sum_{x=0}^1 b(x; 5, 0.2) = 1 - 0.7373 = 0.2627$$

Monalisa Sarma
IIT KHARAGPUR

So, next is another one question. So, what it is given suppose that a system contains a certain type of component whose time in years to failure is given by T . The random variable T so, T is the random variable it is mentioned only its modelled nicely by exponential distribution here our job has become very easy it has already explained that it is exponential distribution

we do not need to worry last question it was not given what distribution is that we need to find out by seeing the problem.

So, here it is already mentioned that it is exponential distribution. So, T is modelled nicely by exponential distribution with mean time to failure that is $\beta = 5$ at a specified β this is one part of the question then it is if 5 of these components are installed in different system, what is the probability that at least 2 are still functioning out of 5 2 are functioning at the end of 8 years. See the last part last sentence if 5 of these components are installed in different system, what is the probability that at least 2 are still functioning at the end of 8 years.

Does not it sound similar to binomial distribution the last portion out of 5 trials 2 successes is not it? Out of 5 trials if there are 2 successes that is what the 5 of these components are installed in different system what is the probability that these 2 are still functionally at the end of 8 years. Now out of 5 2 are still functioning this is a binomial distribution. Now, the binomial distribution means what is the probability of success and binomial distribution to probability of success that means here.

What is the probability success probability that the component is still functioning at the end of 8 years that is my probability of success here probability that the component is working till at the end of 8 years. Now, how do I find a probability that a component is working at the end of 8 years as it is already told that it is exponentially distributed, since it is exponentially distributed, I can just put it in a formula and then I can find out what is this probability?

Probability of $T > 8$ is nothing but putting just simply putting it in a formula of exponential distribution $1 / \beta^{-1} \beta dt$, β is 5 so $1 / \beta e^{-3/5} 1 / \beta$ is not it? dt , so, I got this is my probability. Now, my probability of success means probability that it is functioning at the end of 8 years is 0.2. Now simple binomial distribution at least 2 are still functioning means 2 are functioning, 3 are functioning for $2 + 3 + 4 + 5$ thus instead of doing that.

I can do it $1 - 0$ to 1, the 0 to 1 means does a cumulative from 0 to 1 it is cumulative. So, cumulative means I can see it from the table. So, next one is lognormal distribution.


(Refer Slide Time: 28:06)

Lognormal Distribution

- The lognormal distribution applies in cases where a natural log transformation results in a normal distribution.
- This distribution is widely used in reliability engineering.
- This distribution is defined for positive values of rv (usually time).

Definition: Lognormal Distribution

The continuous random variable x has a lognormal distribution if the random variable $y = \ln(x)$ has a normal distribution with mean μ and standard deviation σ . The resulting density function of x is:

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$


Monalisa Sarma
IIT KHARAGPUR

Lognormal distribution also like exponential distribution and gamma distribution is a widely used distribution. Till now, the distribution that we have studied exponential, gamma and normal are widely used distribution normal mean standard normal basically. So, now lognormal distribution is also has also it is used in reliability studies. So, this mainly it is used to model the repair times not the failure times when you talk a failure times, we do not use lognormal distribution, we use normal distribution or maybe exponential distribution.

But lognormal distribution we use when we try to model the failed repair time of a component. So, and lognormal distribution applies in cases where natural log transformation result in a normal distribution. So, if x is log normally distributed, what the person means, if x is log normally distributed, then log of x is normally distributed, the meaning of the first statement is that if x is log normally distributed, then log of x is normally distributed.

So, this distribution is widely used in reliability known mainly for modelling the repair times, this distribution is defined for positive values of random variables only usually 10. Definitely when we talk about repair time's repentance cannot be negative. So, this distribution is defined for positive values of random variables only. So, this is the probability density function. This is the probability density function for lognormal distribution.

Similarly, we can find a cumulative, if you do the cumulative distribution function that is from $-\infty$ to the particular value x whatever it is, then we see when we can easily use the standard normal distribution table to find out the values are lognormal as well.

(Refer Slide Time: 30:02)

Lognormal Distribution

The mean and standard deviation (SD) of the Lognormal distribution are given by:

$$\mu = e^{\mu + \frac{\sigma^2}{2}} \text{ and } \sigma^2 = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

Monalisa Sarma
 IIT KHARAGPUR

So, this is the value of mean and σ^2 this just what is the simply how we find out the mean? Mean is $\int f(x)$. Variance is $\int x - \mu^2 f(x)$ by solving that if we know by putting that in this value and by solving it, we will be getting this 2 value. Now, there is one problem which you have to solve, before solving this problem, we will this let us see this question how as I told you this lognormal distribution for when we try to find out a cumulative lognormal distribution, we can use the standard normal table we can very well use the standard normal table for that, let us see how we can do that.

(Refer Slide Time: 31:02)

$$F(t) = \int_0^t \frac{1}{t \sigma \sqrt{x}} \exp\left[-\frac{(\ln(x) - \mu)^2}{\sigma^2}\right] dx$$

$$z = \frac{\ln(t) - \mu}{\sigma}$$

$$dz = \frac{dx}{t x}$$

So, what is the lognormal if I write it in terms of cumulative, so, it is F of t is equals to I can write 0 to t F (t) means, cumulative means always it is from - infinity to a particular value. So, here since I am talking of what to say in the lognormal, it is always a positive value. So, I am not a minus infinity I am taking 0. So, 0 to t that is 1 by what is the expression $1 / t \sigma \sqrt{2}$

pi, it is a very big expression, it is difficult to remember so, it is always a better to use the lookup table.

So, it is $\exp(-\ln(t - \mu / \sigma^2))$ and this is whole square dt this was the expression for Ft probability density function. So, now, when we want to cumulative that will be integration from 0 to t. So, this is the expression for cumulate. Now, if I substitute say Z is equals to earlier case I substituted $Z = x - \mu / \sigma^2$ here since, F t is log normally distributed log t is normally distributed.

So, I can substitute $Z = \ln t - \mu / \sigma^2 - \mu$ whole square by saying sorry I will not use the square term here that made up I will not use this square here. So, this is my Z what will be my dz? $dz = dt / \sigma t$ is not it? So, now, substituting this values like how we have done for normal similarly, substituting this $\ln t - \mu / \sigma^2$ if I substituted by Z then what will be and similarly substitute $dt / dz \sigma t$ then what I will get?

I will get the same whatever I got it for the normal distribution so, I will get it in my Ft firstly, let me take a separate space from this.

(Refer Slide Time: 33:30)

The image shows a blackboard with handwritten mathematical work. On the left, there is a sketch of a normal distribution curve. In the center, the cumulative distribution function is written as $F(t) = \frac{1}{\sqrt{2\pi}} \int_0^t \dots dz$. To the right, the substitution $Z = \frac{\ln(t) - \mu}{\sigma}$ is shown, along with the standard normal density function $\phi\left(\frac{\ln(t) - \mu}{\sigma}\right)$.

So, my Ft will be equals to $1 / \sqrt{2\pi}$ after substitution whatever I got it for standard normal similarly, this is that it will be from sorry it is not minus 0 to log of t - μ / σ because here the range is also changes earlier it was from x. Now, what is $x = \log$ of t - μ / σ I have substituted that so, it becomes $\exp - Z^2 / 2 dz$ this is the same thing, but we got it when we have done for standard normal distribution instead of this value here we had what here we had x.

Now, x or t whatever it is now, here we are having \log of $t - \mu / \sigma$ that is all. So, now, for this we have referred to the table that is standard lookup table at what point you have to refer to the table. Table usually we call this value this is the symbol we use earlier case we have referred to this value it is Φ of x means at portion x is not it? If this was the normal distribution suppose, this is my x , this is Φ of x this gives me this value now, here so, what I will find out from that normal table?

I will find a Φ this \log of $t - \mu / \sigma$ that is all, that will give me my value of this area probability of this area from 0 to less whatever it is solving this problem will be clear.

(Refer Slide Time: 35:10)

Lognormal Distribution—Example 5

Problem

Concentrations of pollutants produced by chemical plants historically are known to exhibit behavior that resembles a lognormal distribution. This is important when one considers issues regarding compliance with government regulations. Suppose it is assumed that the concentration of a certain pollutant, in parts per million, has a lognormal distribution with parameters $\mu = 3.2$ and $\sigma = 1$. What is the probability that the concentration exceeds 8 parts per million?

Solution

Let random variable X denotes pollutant concentration $P(X > 8) = 1 - P(X \leq 8)$
 Given, $\ln(X)$ has a normal distribution with $\mu = 3.2$ and $\sigma = 1$

$$P(X \leq 8) = \Phi\left(\frac{\ln(8) - 3.2}{1}\right) = \Phi(-1.12) = 0.1314$$

Here, Φ = the cumulative distribution function of standard normal distribution.

So, concentration of pollutants produced by chemical plants historically are known to exhibit behaviour that resembles a lognormal distribution, this is information for us concentration of pollution produced by chemical plant it resembles a lognormal distribution. This is important when one considers issues regarding compliance with government regulation and other information. Suppose it is assumed that the concentration of a certain pollutant in parts per million has a lognormal distribution with parameter $\mu = 3.2$, $\sigma = 1$.

So, it is given what it is given that the concentration of the certain pollutant in parts per million of course, it has a lognormal distribution with this is these are the parameters μ and σ . So, what is the probability that a concentration exceeds 8 parts per million. So, that is basically we have to find a probability of x greater equal to 8. And problem is greater equals

to 8 means, if I say this is if I convert if my x is a lognormal distribution, my log of x will be normal distribution. So, what I need to find out?

So, first this I will have to convert it to standard normal variable how $\log x - \mu / \sigma$. And basically, I need to find a, it is greater than what value, we will be getting this value suppose my Z value is this $\log x - \mu / \sigma$, this is whatever is Z value. So, this is Z value from the table I can get this value. But what I needed is greater than this, I need this portion. So, it will be 1 minus of this I will be getting the value. See here P of X greater than 8 this now, simple put just finding out the Z value.

But I told you we just find out the Z value \ln of 8 - 3.2. Because why we are finding out is Z value? Because integration is from 0 to this Z value. What was the Z value? \ln of $x = \mu / \sigma$ and this whatever expression that is probability density function and that we do not have to calculate it because we know what the value corresponding to this just value, this probability, I am in the area for this value that we can get it from the lookup table. So, this we got Z value we got -1.12.

So, again from the normal distribution table, what is the value corresponding to 1.12 is this, we are interested in finding out greater than 8. So, 1 minus this will give us that value that is the value greater than 8. So, now, next distribution, that is there is one important distinction that is that Weibull distribution, but the Weibull distribution, I will not be discussing in this class, though this Weibull distribution, I have not included in the course also.

But I want I would like to mention that it is a very important discussion for reliability studies. But because this course audiences not only students who are from reliability, maybe none of them are from reliability background, you do not have any knowledge on reliability. So, Weibull distribution will be very a bit complicated for them to understand it is a bit high standard. So, what I am doing is that people who are interested I have kept the slides here, people who are interested you please go through this Weibull distribution.

And if you have any doubt we can clear in a doubt clearing session, but I will not be included in this course. That means when I am telling I am not including in this course you will not be you cannot expect an equation from Weibull distribution in your exam as well plus I will not be discussing here also, but if you have any doubt we can if anyone is interested in going

through Weibull distribution, then we can discuss it then any doubt we can discuss in the doubt clearing session.

(Refer Slide Time: 39:02)

CONCLUSION

- In this lecture we learned about some important continuous distribution function that includes the knowledge of –
 - Gamma distribution
 - Exponential distribution
 - Lognormal distribution
- Some solved problems are provided. Learners are requested to solve the practice problems
- In the next lecture, we will cover a tutorial.

Monalisa Sarma
IIT KHARAGPUR

So, now, I end this lecture on continuous probability distribution in this class we have discussed gamma distribution, exponential distribution and lognormal distribution. Of course, we have solved your problem also. And in the next class we will cover a tutorial.

(Refer Slide Time: 39:18)

REFERENCES

- Sheldon Ross , A First Course in Probability (fourth ed.), Macmillan College Publishing, New York (2013)
- D. P. Bertsekas and J. N. Tsitsiklis, Introduction to Probability. Nashua, NH: Athena Scientific, 2008
- W. Feller, An Introduction to Probability Theory and Its Applications, vol. 2. New York: Wiley, 1971

Monalisa Sarma
IIT KHARAGPUR

With this and I here I specify the references and thank you guys.