

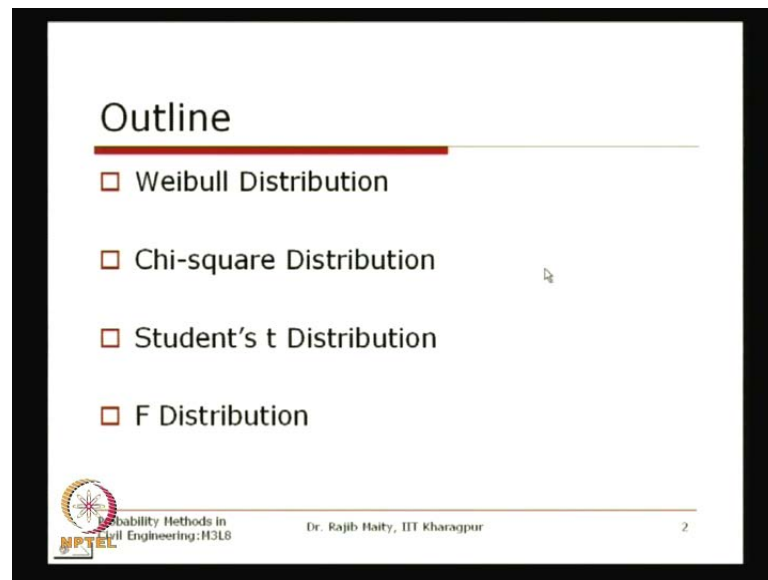
Probability Methods in Civil Engineering
Prof. Dr. Rajib Maity
Department of Civil Engineering
Indian Institute of Technology Kharagpur

Lecture No. #13
Probability Distribution of Continuous RVs(Contd...)

Welcome to this lecture 8 of this module three. This is the last lecture for this module, where we are discussing about random variable their properties, and some standard distribution for continuous random variables as well as discrete random variables. So, up to the last class we discussed some of the continuous random variable, and we discuss that there are few more is pending, but still whatever is covered, and whatever will be covered in this lecture may not be the full list of the standard probability distribution, there could be some other as well, but these are the distribution which are mostly used for different problems in civil engineering.


So, in the last class where the Weibull distribution was pending. So, we will start with the distribution this class, and we will also discuss some of the distribution after that, that we mentioned in the last class that chi square distribution, t distribution, and F distribution which are mainly used for the different statistical test. So, our this lecture on that probability distributions of continuous random variable, we will continue in this lecture as well.

(Refer Slide Time: 01:39)



Outline

- Weibull Distribution
- Chi-square Distribution
- Student's t Distribution
- F Distribution

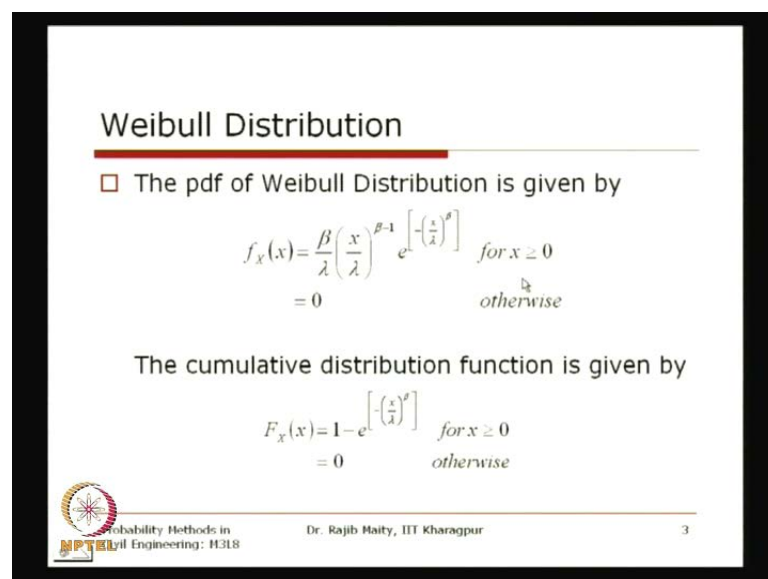
 Probability Methods in
Civil Engineering: H3L8

Dr. Rajib Maity, IIT Kharagpur

2

And we will discuss Weibull distribution first, and then some of the distributions which are specifically important for the different statistical test; and this statistical test will be carried out in last module- module 7, where the properties of this distribution particularly this three distribution that chi square distribution, student's t distribution, and F distribution will be used for different statistical test, that we will discuss. In this lecture, we will just discuss about their basic properties, and which distribution is suitable for what type of test that we will discuss in this class.

(Refer Slide Time: 02:29)




Weibull Distribution

- The pdf of Weibull Distribution is given by

$$f_X(x) = \begin{cases} \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^\beta} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The cumulative distribution function is given by

$$F_X(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\lambda}\right)^\beta} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

 Probability Methods in
Civil Engineering: H3L8

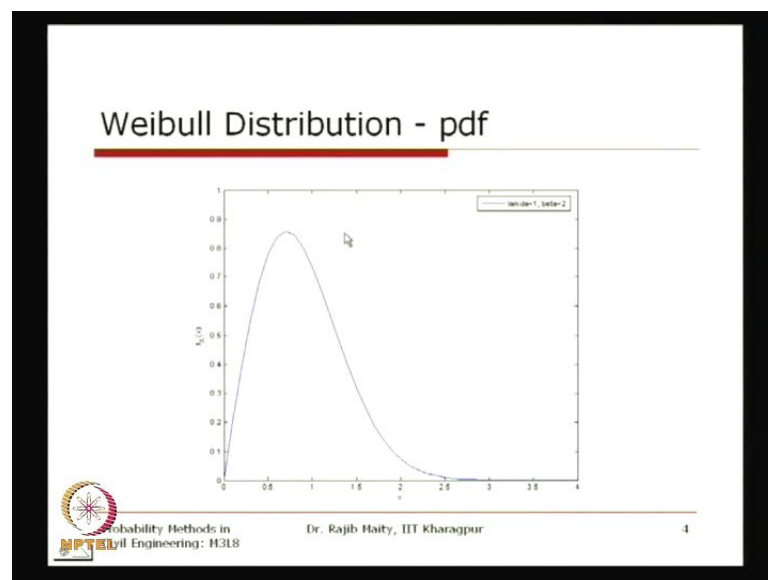
Dr. Rajib Maity, IIT Kharagpur

3

So, we will start with this weibull distribution. This general form of the p d f of weibull distribution looks like this. There are two parameters: one is beta and another one is lambda. So, this p d f probability density function is $\beta \lambda^\beta x^{\beta-1} e^{-\lambda x^\beta}$ for $x \geq 0$.

So, for this support of this distribution is bounded at 0. So, lower bound is 0 and upper bound is infinity. And for the negative values of x , this is not defined; this is 0. And we know that the cumulative distribution function for any p d f is the integration from this lower support that is here is 0 up to here is 0 up to that particular value x . If you do that integration, we get the cumulative distribution function, and this can be shown that, this is $1 - e^{-\lambda x^\beta}$ for $x \geq 0$. So, if we compare this distribution from whatever the distribution we have discussed earlier, then we will see that, we discussed earlier the exponential distribution, we discussed earlier the gamma distribution, and this particular distribution is having some kind of sense of the generalization of those distributions, and we will show that how this distribution is related to those distributions as well, that is exponential distribution and gamma distribution.

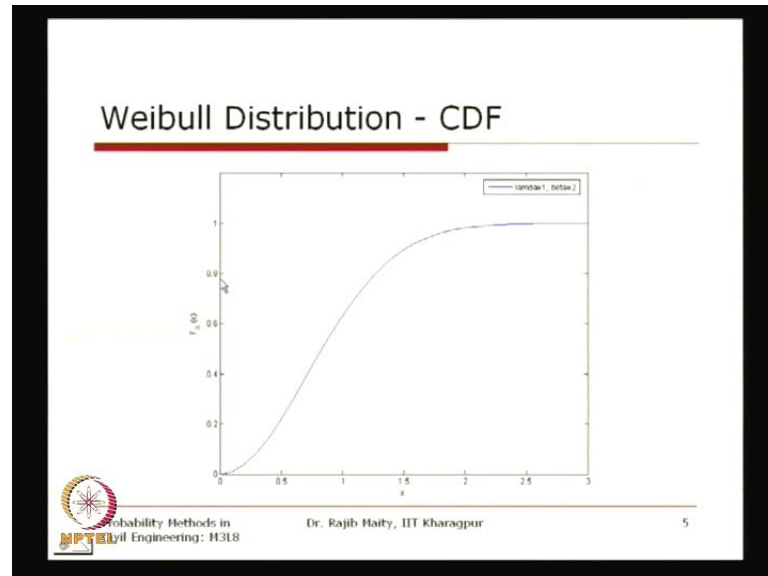
(Refer Slide Time: 04:18)



So, the other properties of this distribution means first this is one example of that weibull distribution p d f, how the p d f look like here the parameters that is taken as this lambda equals to 1 and beta equals to 2. So, taking this to this set of parameter, this is the general shape of this one, which we can see that this is a positively skewed distribution and lower

bound is 0 and upper bound this is coming and asymptotic to 0. So, this is your x and this is your $f(x)$ that is density of probability.

(Refer Slide Time: 04:59)



And for the same combination of the parameters that is lambda equals to 1 and beta equals to 2 that is same $p.d.f$ that is shown in last slide if we take that cumulative distribution that is $F(x)$ is capital $F(x)$, then we will get this type of distribution which is starting from 0 and asymptotic to 1 like this type of distribution we can see.

(Refer Slide Time: 05:27)

Mean, Variance and Coefficient of Skewness of Weibull Distribution

- Mean is given by

$$\mu = \lambda \Gamma(1 + \beta^{-1})$$
- Variance is given by

$$\sigma^2 = \lambda^2 [\Gamma(1 + 2\beta^{-1}) - \Gamma^2(1 + \beta^{-1})]$$
- Coefficient of Skewness is given by

$$\gamma = \frac{2\Gamma^3(1 + \beta^{-1}) - 3\Gamma(1 + \beta^{-1})\Gamma(1 + 2\beta^{-1}) + \Gamma(1 + 3\beta^{-1})}{[\Gamma(1 + 2\beta^{-1}) - \Gamma^2(1 + \beta^{-1})]^{3/2}}$$

The slide footer includes a logo, 'Probability Methods in Mechanical Engineering: M318', 'Dr. Rajib Hait, IIT Kharagpur', and the slide number '6'.

So, now, this first few moments if we can discuss about this one that this mean is given by $\mu = \lambda \Gamma(1 + \beta^{-1})$. This is one parameter and this gamma function of $1 + \beta^{-1}$ is $1 + 1/\beta$. We discuss about this gamma function, some previous lectures while discussing that beta distribution and gamma distribution. So, if we know the value of this β we will be able to know what is the value of this gamma function and we will be able to calculate this μ . So, this μ is dependent on both this parameter we have seen and similarly the variance also is given by this expression which is $\lambda^2 \Gamma(1 + 2\beta^{-1}) - \mu^2$ and $\Gamma(1 + \beta^{-1})^2$.

Basically this one if you just this λ^2 if you take inside this bracket this becomes that mean square. So, that mean square. So, we can say this quantity that is the λ^2 multiplied by $\Gamma(1 + 2\beta^{-1}) - \mu^2$ will be the variance of the distribution and similarly, if we take that skewness we will get along expressions like this which is the measure of the skewness of the Weibull distribution.

One thing is important here that, if we discuss that if we simply prove that $\beta = 1$ here then, you know that this will become that $\Gamma(1 + \beta^{-1})$.

(Refer Slide Time: 07:14)

Handwritten derivation on a grid background:

$$\begin{aligned} \beta &= 1 \\ \mu &= \lambda \Gamma(1 + \beta^{-1}) = \lambda \Gamma(1 + 1) = \lambda \Gamma(2) \\ &= \lambda \\ \Gamma(2) &= 1 \Gamma(1) = 1 \\ \sigma^2 &= \lambda^2 [\Gamma(1 + 2\beta^{-1}) - \Gamma^2(1 + \beta^{-1})] \\ &= \lambda^2 [\Gamma(3) - \Gamma^2(2)] \\ &= \lambda^2 [2 - 1^2] \\ &= \lambda^2 \end{aligned}$$

Side calculation:

$$\begin{aligned} \Gamma(3) &= 2 \Gamma(2) \\ &= 2 \end{aligned}$$

So, what we are trying to say here if we just put that $\beta = 1$ then that μ that $\lambda \Gamma(1 + \beta^{-1})$ which is equal to $\lambda \Gamma(1 + 1)$ is $\lambda \Gamma(2)$. So, and from the earlier lectures we know this $\Gamma(2)$ is 1.

integer then, we know that this if gamma equals to n minus 1 gamma minus 1. So, this is 1 gamma 1 and this gamma 1 value is equals to 1. So, this is basically 1. So, this expression that is mu becomes lambda. Similarly, if I take that sigma square which is the expression is lambda square multiplied by gamma 1 plus 2 beta inverse minus gamma square 1 plus beta inverse. So, putting this beta equals to 1 it will eventually come that lambda, this lambda square gamma this is three minus gamma square 2.

So, this is 1 and this will become your 2. So, lambda square this is 2. So, gamma 3 you know gamma 3 equals to 2 gamma 2 and gamma 2 equals to 1. So, it is 2 minus 1 square. So, this is equals to lambda square if we see that for the exponential distribution we got exactly the same value of this moments at least the first moments that is for an exponential distribution; that means, is equals to lambda and variance is equals to lambda square. So, if we put this beta equals to 1 it seems that it is, it is becoming an exponential distribution.

(Refer Slide Time: 09:14)

Weibull Distribution

□ The pdf of Weibull Distribution is given by


$$f_X(x) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda} \right)^{\beta-1} e^{-\left(\frac{x}{\lambda} \right)^\beta} \quad \text{for } x \geq 0$$

$$= 0 \quad \text{otherwise}$$

The cumulative distribution function is given by

$$F_X(x) = 1 - e^{-\left(\frac{x}{\lambda} \right)^\beta} \quad \text{for } x \geq 0$$

$$= 0 \quad \text{otherwise}$$



Probability Methods in
Civil Engineering: M318

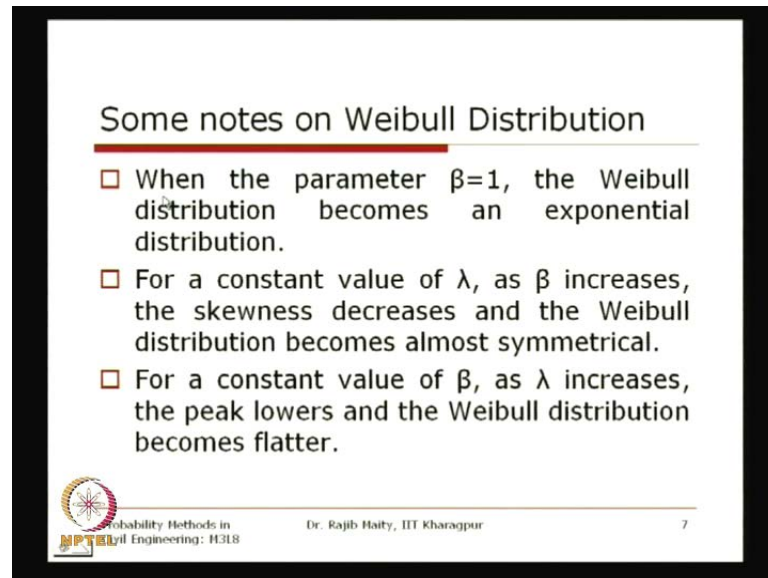
Dr. Rajib Maity, IIT Kharagpur

3

So, if we just go back to the distribution of this Weibull distribution and here if you put that beta equals to 1 then we can see that that is the 1 by lambda and here it is beta equals to 1. So, this is 1 and this is power minus x by lambda for x greater than 0 and otherwise it is 0. So, which is nothing but exactly same to this exponential distribution, exponential distribution. So, here with respect to the exponential distribution we can say that one more parameter is introduced which is beta.

If this beta is equal to 1, this Weibull distribution becomes an exponential distribution and we will also show that how this Weibull distribution is also that gamma distribution is a special case for this Weibull distribution that we will discuss in a minute. So, now, this we discuss and this first few moments mean variance and coefficient of skewness we discuss.

(Refer Slide Time: 10:14)



Some notes on Weibull Distribution

- When the parameter $\beta=1$, the Weibull distribution becomes an exponential distribution.
- For a constant value of λ , as β increases, the skewness decreases and the Weibull distribution becomes almost symmetrical.
- For a constant value of β , as λ increases, the peak lowers and the Weibull distribution becomes flatter.

Probability Methods in
Civil Engineering: H318

Dr. Rajib Maity, IIT Kharagpur

7

Some note that what just now, we discuss is the first note that when parameter beta equals to 1 the Weibull distribution becomes an exponential distribution which just now we discuss.

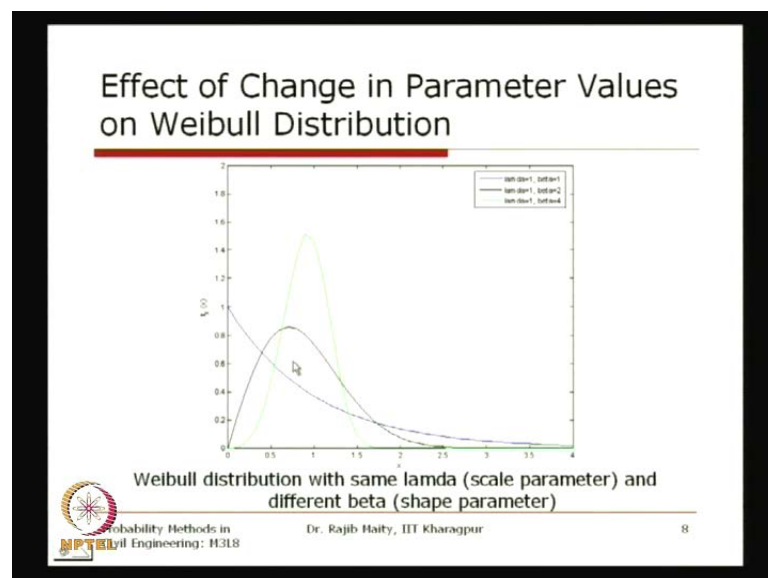
Now, if we just want to see the effect of the parameters that is both for this lambda and beta. So, if we just keep one parameter constant and if we just increase the other parameter then, what will be the effect on the shape of this p d f of Weibull distribution?

So, the first it says that for a constant value of lambda. So, if you just take constant value of lambda and if we increase this beta. So, as beta increases the skewness of, the skewness decreases and the Weibull distribution becomes almost symmetrical. So, this is important in the sense that when we need to model some kind of, some random variable which is asymmetrical as well as this is lower bounded by 0. Then, we can say, we can say that this Weibull distribution maybe a suitable choice for that case because we can

estimate this beta is high and then this weibull distribution becomes symmetrical. That we will see that, how it is changing with best on the different values of this beta.

On the other hand, if we take this beta as constant if we take some constant value of beta and if we increase this lambda, then the peak lowers and the weibull distribution becomes more flatter. So, weibull distribution becomes more flatter means that, that its peakedness reduces, its skewness also reduces, it becomes more flatter. So, if we see that its shape depending on this different combination of this parameter that is alpha and beta then this will be more clear to us.

(Refer Slide Time: 12:25)



So, this is the first thing when we are keeping this lambda constant and adding this beta and beta equals to 1 2 and 4 and keeping the all main all three cases is lambda is equals to 1. What you can see here that this for this blue line this pdf shown by blue line this is for the combination that lambda equals to 1 and beta equals to 1 for this black one, the combination is the lambda equals to 1 and beta is equals to 2 and for this green one this lambda equals to 1 means all the lambdas are same for all three and beta equals to 4.

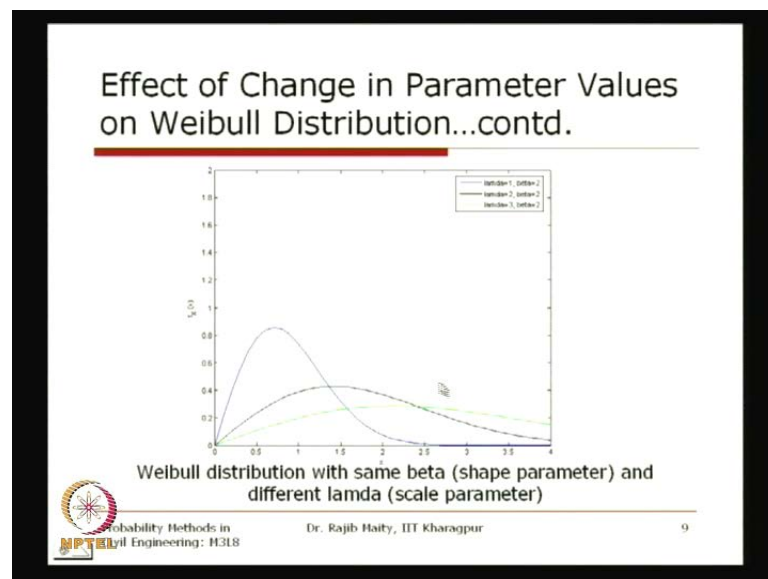
So, from this blue then black, then green the beta is increasing from 1 2 and then 4. So, here 2 things we can see, the first is that just now we discussed that if beta is equals to 1 when we say that beta is equals to 1; that means, we are talking that irrespective of any value of lambda. So, when beta is equals to 1 it approaches to the exponential

distribution. So, that we can see herethat, this is anthis an exponential distribution and we know that this at 0 at x is equals to 0 the value of f_x is 1 by λ .

So, as we are taking the λ is equals to 1 that is why that x is equal to 0 it is coming to be 1. Similarly, now when we are increasing this β from 1 2 and 4, we can see that this is becoming some, this is these are positively skewed we know exponential distribution is positively skewed this can be seen graphically as well as mathematically also we've discussed in previous lectures. Now, when we increase β from β equals to 1 to 2 the skewness that asymmetry decreases.

So, it approaching to the symmetrical and it is more clear when we are going to this four that it is becoming some more to more symmetrical compared to the other two graphs where β equals to 1 and β equals to 2. So, if we further increase β we can see that it is, it is becoming more symmetrical with respect to its mean. Mathematically also this can be this can be shown that we will see in a minute.

(Refer Slide Time: 14:59)



Second case if we just keep this β constant and if we, if we change the value of λ here, we have shown this three different PDFs with different combination of λ and β . So, this first one this blue one is for this λ equals to 1, β equals to 2; this black one is the λ equals to 2, β equals to 2 and this green one is for λ equals to 3 and β equals to 3. Thus for all three PDFs here the β is same which is equal to 2 and for the, and the λ is increasing from 1 then for the black it is


2 and for green from it is 3. So, what we can see is that when this lambda is low, then its peak is more and with the increase of this lambda this peak becomes flatter.

Basically, when we are having some set of data there are some tests to understand that which distribution it is following and after we select a particular distribution, we generally discuss that we generally want to know, what is its parameter?, then this parameter is estimated from the sample data that is available to us that is basically the first step for any probability model, whatever the problem is faced. So, the sample data is collected from the sample data first we test that which distribution this is following and then we estimate the parameter. So, that while estimating the parameter there are different methods available. So, here what we are discussing is the effect of this parameter the graphical significance of this parameter. If we select one parameter, if we change one parameter how the shape and location of particular p.d.f changes, that is what we are discussing for all the, all the models, all the probability distributions.

(Refer Slide Time: 17:02)

Some notes on Weibull Distribution...contd.

- When the Weibull distribution is used to model the failure time of a system, the shape parameter β has the following interpretations:
 - If $\beta > 1$, it indicates that the failure rate is increasing with time i.e., the tendency to fail is increasing with aging of the system
 - If $\beta = 1$, it indicates that the system has a constant failure rate
 - If $\beta < 1$, it indicates that the failure rate is high initially and decreases gradually as the defective elements are removed from the population



Probability Methods in
 Reliability Engineering: H318

Dr. Rajib Haitry, IIT Kharagpur

10

Now, one important discussion with respect to this Weibull discussion is this the parameter beta, this parameter beta is playing a vital role and it is having different meaning for the different ranges of this parameter first. So, when the Weibull distribution is used to model the failure time in a system the shape parameter beta has the following interpretations.

So, how this beta we can interpret for the different values or the different ranges of this parameter beta. So, first if we start with this beta equals to 1 that is the, when we say

that beta equals to 1. We know that this is the approaching to the exponential distribution. Now the failure of a system; that means, that we know that value is constant that is that failure rate, if we while discussing that exponential distribution we discuss that this the failure rate or the interarrival rate, average time on which a particular phenomenon is occurring that is constant overtime based on that assumption, we develop that exponential distribution this is exactly what is stated here that is $\beta = 1$ it indicates that the system has a constant failure rate; that means, which is the basic assumption of the exponential distribution and we have also discuss that when beta equals to 1 the distribution is approaching to the exponential distribution. Now there are two other region if beta is less than 1 or if beta is greater than 1, what does it implies?.

So, when we say that this beta is less than 1, it indicates that the failure rate is increasing with time. That is the tendency to fail is increasing with the aging of the system, this is important for different civil engineering problem for the, for any structure that we take due to its aging problem or some. Some apparatus for its fatigue problem when the time increases sometimes the inter arrival rate of the phenomenon may increase.

So, here we are deviating from the basic assumption of this exponential distribution that this rate of this failure need not be constant over the time starting from 0 to the any positive infinite towards the plus infinity direction. So, when we say that this rate of failure is, rate of failure is increasing with time; that means, that is captured when the distribution is having the parameter that having the particular beta is greater than 1.

Similarly, So, this is you know that when some particular, some particular structural member or the structure itself due to its aging problem or when it is passing more time the chances of failure is increases. So, that kind of case this weibull distribution with beta greater than 1 will be the, will be a suitable case. On the other hand when the beta is less than 1 it indicates that the failure rate is high initially and decreases gradually as the defective elements are removed from the population. Now, rather to come to this civil engineering problem, if we just take a social problem then it will be a more clear that, if we say that due to the health means health condition due to the this is the health system. If the mortality of the babies is increase at birth then this is a one of this, one of this very suitable application to model that kind of situation with this beta less than equals to 1.

So, at the birth is shown the mortality rate is more compared to when the age increases of this of the child. Similarly, for if we come to this some application of civil engineering


that time when we are producing some kind of product and we are just marketing it. Now, if there is any defective element in the population then what happens and if there is any the initial check of this whatever the product is developed then there is a chance that, if we can separate out those defective materials then, that can either fail at the initial stage itself or if we pass from this initial stage then it can go for a long time. So, **what we** what I am trying to communicate is that this rate of failure at the initial time is more compared to that when it increases. So, at that case that beta is, beta is less than 1. So, that the failure rate is high initially and decreases gradually as time passes.

So, these are the interpretation of this parameter beta for a weibull distribution and suitably we can think of that for what kind of situation which range of beta is suitable for different problems in hand.

(Refer Slide Time: 22:35)

Examples of Weibull Distribution in Civil Engineering

- ☐ Weibull distribution is often used to study failure strengths of engineering materials and also to describe the failure time of machines.
- ☐ Probability distribution of wind speed is sometimes modeled as Weibull distribution.
- ☐ Hydrological variables, such as, annual maximum daily rainfall depth can be modeled by Weibull distribution.



Probability Methods in
Civil Engineering: H318

Dr. Rajib Maity, IIT Kharagpur

11

These examples of the weibull distribution in civil engineering as we discussed that this weibull distribution is often used, to study the failure strength of engineering materials and also to describe the failure time of the machine. So, it can. So, the failure time of a machine or the failure strength of engineering material. If I assumed to be its to be constant then it can be as well an exponential distribution; that means, weibull distribution with parameter beta equals to 1 or it depends on that, what kind of situation? what kind of problem? in hand based on that the beta parameters may change.

Secondly, sometimes the probability distribution of the wind speed particularly the hourly wind speed is also seen to follow the weibull distribution. So, that is also designed using the weibull distribution, in the application of water resource and hydrology the annual maximum daily rainfall depth, if I take up to that data also has found to follow a weibull distribution with suitable combination of its parameters λ and β .


(Refer Slide Time: 23:57)

Chi-square Distribution

□ If Z_i ($i=1, 2, \dots, n$) are n independent standard normal random variables, then the sum of the squares of n random values of Z follows a Chi-square Distribution with n degrees of freedom.

Thus,
$$X = \sum_{i=1}^n Z_i^2$$

where $Z = \frac{Y - \mu}{\sigma}$ and $Y \sim N(\mu, \sigma)$



Reliability Methods in
Civil Engineering: H3L8

Dr. Rajib Maity, IIT Kharagpur

12

Next, we will discuss three distributions one after another, that we will start with the chi square distribution and these distributions are mostly as I mentioned earlier these distributions are mainly important for the statistical test and different types of test is carried out using different types of distributions. So, first we are discussing this chi-square distribution which is generally used for the purpose, or for the purpose of goodness of fit. So, many times I have mentioned that once the sample data is available to us. We generally have to first we have to test that distribution it is following and. So, this is also can be tested statistically and that test is generally done through this chi square test it is called and based on this chi square distribution. So, we will understand now the what are this distribution and what are the different properties.

So, first we are starting this chi square distribution, if Z_i equals to 1 to n that is Z_1, Z_2, Z_3 up to Z_n are independent and identically distributed of course, we use the word this

3,3 were together independent and identically distributed sometime it is abbreviated as iid here these are independent of course and following the standard normal distribution; that means, these are identically distributed. So, independent standard normal random variables then their sum of the squares. So, this is important the square the sum of the squares of this n random variables, that follow a chi-square distribution with n degrees of freedom. So, here that x that the random variable that I am defining is equals to the summation of those Z 1 square plus Z 2 square plus up to Z n square. So, where this Z are the standard normal distribution.

So, if I say that this Y is following a normal distribution with mu and sigma their parameter. Then we can make it, we can get their reduced variate Z which is equals to Y minus mu by sigma and this Z for this n numbers of independent random variables are their. If we square them of add them then these random variable that summation x that is follow a chi square distribution.

(Refer Slide Time: 26:42)

Chi-square Distribution...contd.

- The pdf of Chi-square distribution is given by


$$f_{\chi^2}(x) = \frac{x^{\frac{(v-2)}{2}}}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} \exp\left(-\frac{x}{2}\right) \quad x \geq 0$$

$$= 0 \quad \text{elsewhere}$$

where v is a positive integer known as degree of freedom
- The cumulative distribution function is given by

$$F_{\chi^2}(x) = \int_0^x \frac{x^{\frac{(v-2)}{2}}}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} \exp\left(-\frac{x}{2}\right) dx \quad x \geq 0$$

$$= 0 \quad \text{elsewhere}$$



Reliability Methods in
Civil Engineering: M3L8

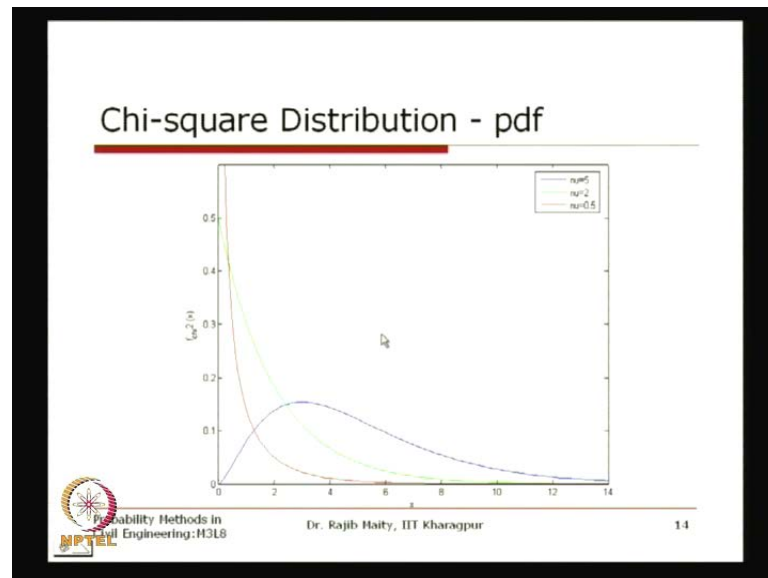
Dr. Rajib Maity, IIT Kharagpur

13

Now the p d f of this chi square distribution is like this that is, this is f chi square x is x power nu minus 2 by 2 divided by 2 power nu by 2 gamma nu by 2 exponential minus x by 2 for x greater than equals to 0. Here nu is the parameter known as the degrees of freedom this is an integer it is a positive integer and this distribution is again lower bound by 0 and for negative values it is 0. So, this nu here mentioned that this nu is a

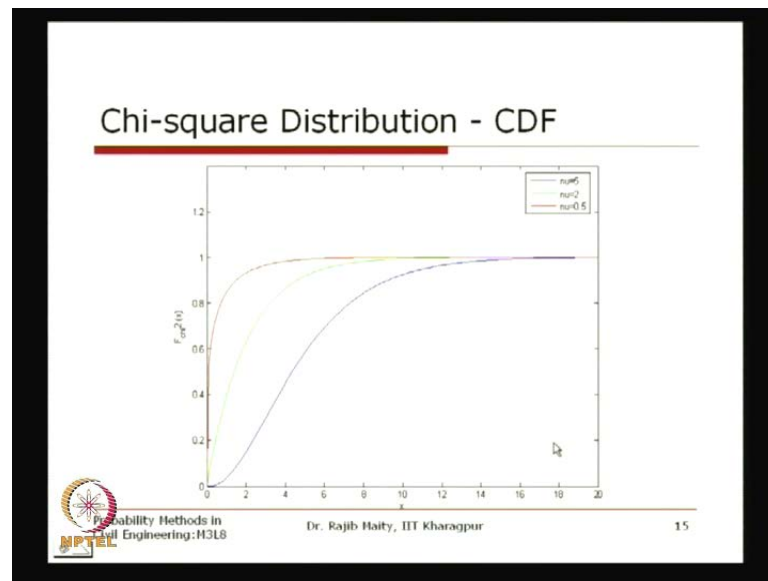
positive integer known as degree of freedom. Again, we know that from this distribution this cumulative distribution is nothing, but the integration from the lower support to a specific value x of the p d f and this is also defined for this x greater than equal to 0 and 0 elsewhere.

(Refer Slide Time: 27:49)



So, we will know that its different moments and different initial moments. Now, before that we will show how this for this different values of this ν which is that the degrees of freedom for this chi square distribution is this. So, here we have used different combination of this ν and we are. So, showing that as this ν increases, then this is becoming this. So, this blue is for the maximum values of ν compared to this green and this red here. So, this is if its if the degrees of freedom increases is generally become more symmetrical.

(Refer Slide Time: 28:39)



And for the and these are the for those three chi square distribution, these are the cumulative distribution for those pdf that is shown in this last slide.

(Refer Slide Time: 28:57)


Mean, Variance and Coefficient of Skewness of Chi-square Distribution

- Mean is given by

$$E(\chi^2) = \nu$$
- Variance is given by

$$Var(\chi^2) = 2\nu$$
- Coefficient of Skewness is given by

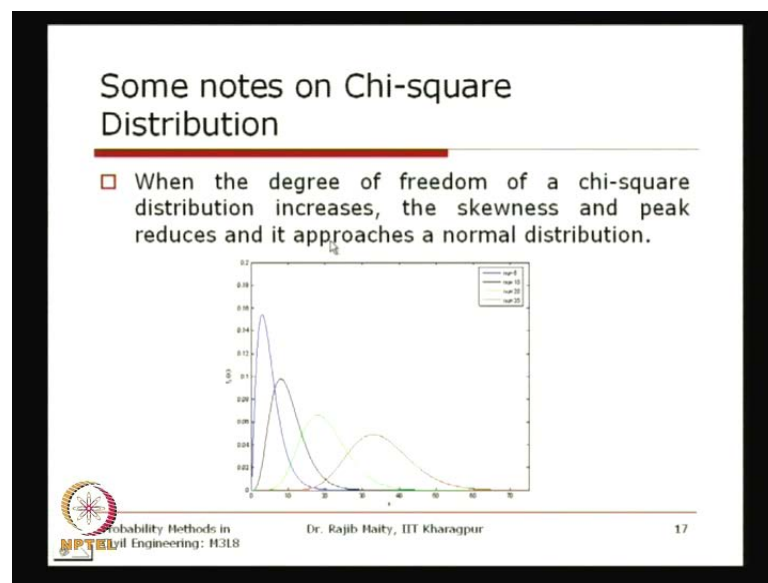
$$\gamma_{\chi^2} = \frac{2}{\sqrt{\nu/2}}$$


 Probability Methods in
Civil Engineering: M3L8
 Dr. Rajib Haity, IIT Kharagpur
16

Now, the initial few moments if we see then there are some interpretation particularly its skewness as just we have seen from this graph the first is this its mean. The mean of this chi square distribution is equal to its degree of freedom which is ν the variance of the

chi-square distribution is $2/\sqrt{\nu}$ that is double of the mean and the coefficient of skewness if we see then, this is that $2/\sqrt{\nu}$ by $\sqrt{2}$. So, this $2/\sqrt{\nu}$ by $\sqrt{2}$. So, this is always positive. So, the chi square distribution always positively skewed distribution and you can see here what we have seen graphically in a minute back. Also, that if we increase this ν that is if the degree of freedom is increased. Then $2/\sqrt{\nu}$ then this becomes lower and lower and for every very high degree of freedom, this becomes almost symmetrical as this skewness is reducing with the increase of this ν .

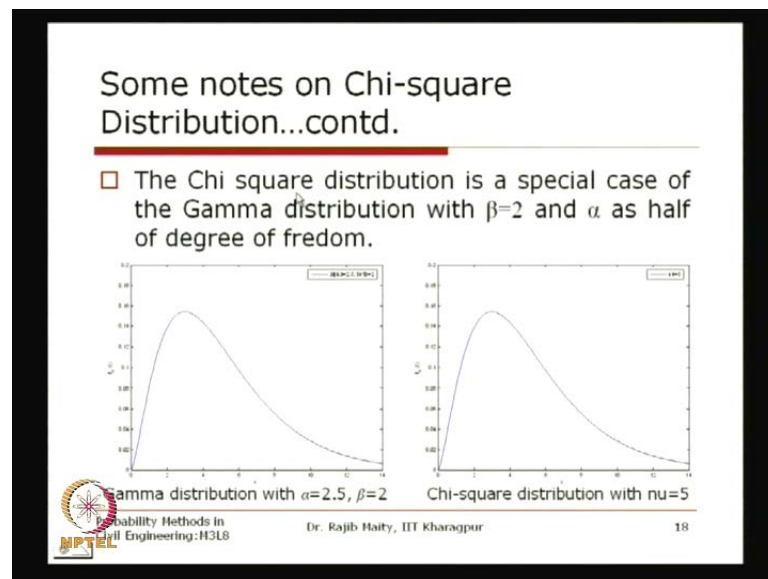
(Refer Slide Time: 30:12)



So, this is again explain here graphically that is when the degree of freedom of a chi square distribution increases the skewness and peak reduces, and it approaches to a normal distribution in the sense that it is approaching to a symmetrical distribution and going away from its origin. Here we have taken the four examples with different value of degree of freedom. So, the first one the blue one that we see that is for the degrees of freedom that is ν equals to 5, the black one that ν equals to 10 then this green one is 20 and this ν is 35.

So, here you can see that as this degrees of freedom is increasing this distribution becomes almost symmetrical of course, all these distributions are lower bounded by 0 that is the basic property of this chi square distribution but as this one increases this approaches to a approximate normal distribution with high value of this degrees of freedom.

(Refer Slide Time: 31:30)




Now, this chi square distribution is a special case of the gamma distribution with β equals to 2 and this α as the half of the degree of freedom. So, this is there are two graphs are shown is that, this is the, this is this if you see this is chi square distribution with ν that that is the ν the parameter of this chi square distribution is equals to 5.

So, if we take this one that ν equals to 5 then the shape look like this the p d f shape look like this now, if we this is a gamma distribution with two parameters we know that this gamma distribution is having two parameters. The first that β is 2 and α is equals to half of the these degrees of freedom that is 2.5. So, if we plot a gamma distribution with α equals to 2.5 then β equals to 2, then we get the exactly the same p d f with for which is for the chi square distribution with degrees of freedom 5. So, this is the relation between this chi square distribution and the gamma distribution.

(Refer Slide Time: 32:50)

Some notes on Chi-square Distribution...contd.

- If X_i be a series of independent, chi-square distributed RVs with parameters ν_i , then their sum $X = \sum X_i$ is also a chi-square distributed RV with parameter $\nu = \sum \nu_i$.
- If a number of random samples X_1, X_2, \dots, X_n are independent and identically distributed RVs with $N(\mu, \sigma^2)$, then the RV $\sum_{i=1}^n (X_i - \bar{X})^2$ has a chi-square distribution with parameter $\nu = n - 1$.
- The RV $(n-1)s^2/\sigma^2$ is also a chi-square distributed RV with parameter $\nu = n - 1$.



Reliability Methods in
 Civil Engineering: M3L8

Dr. Rajib Haitry, IIT Kharagpur

19

Now, if that x_i be a series of independent chi square distribution, chi square distributed random variables with parameter ν_i that is for first that X_1 this degrees of freedom that the parameter is ν_1 X_2 parameter is ν_2 then their sum, if I take the sum of those all.

All those X_i 's that is from 1 to n then the a new random, will get a new random variable X which is the summation of them simple arithmetic summation then this new random variable is also a chi square distribution with the parameter is equals to summation of all those degrees of freedom. So, this is the additive rule you can say for this chi square distribution that is if we add more than one chi square distribution and if we know they are what is the degrees of freedom. Then we can get a new random variable which is also chi square distribution and degree of freedom for this new random variable is equals to the summation of the degrees of freedom for all those random of, all those random variables we added up. So, that is explained here for this first note of this chi square distribution.

The second is the, if a number of random variable that is X_1, X_2, \dots, X_n are independent and identically distributed random variables with normal distribution, with the parameters μ and variance σ^2 mean μ and variance σ^2 then, the random variable if we take here that is $\sum_{i=1}^n (X_i - \bar{X})^2$ then, sum them up these variable this is

also a random variable which will follow a chi square distribution with parameter that ν is equal to $n - 1$.

Now, there are two things here the first one is that we have n random variables all are, all are random all are normal distribution that is identical distribution which is normal distribution with their parameters. Now, if we get the another parameter we know that if we just make them square this is what we started the discussion of our chi square distribution earlier, but here what we are doing is that we are instead of calling this random variable to the standard normal distribution, we are just making it general that is it these are the normal distribution with some parameter of mean and variance. Then, if we add them up after deducting their mean this \bar{X} is the sample mean. So, this \bar{X} is equal to the summation of all their values divided by the n . So, that is that average of, that is the mean that is the sample estimate of the mean. So, if we take out that mean and then square them up and then add them **of**, then we will also have a new random variable which is; obviously, follow the chi-square distribution, but earlier we told that this follow a chi-square distribution with parameter that ν equals to n because we are adding n degrees of freedom, but here the degrees of freedom is that $n - 1$. So, this one degree of freedom is lost and this loss of this degrees of freedom is due to the calculation of their mean. Because, I have to calculate their mean from the sample that I have to deduct from them to and then square it up. So, while calculating this mean the one degree of freedom is lost and that is why this, degrees of freedom for this random variable which is chi square distribution this degrees of freedom is $n - 1$.

Another random variable if I take that this s , s is now the sample estimate of, sample estimate of the standard deviation. So, this s^2 divided by this σ^2 , σ^2 is this population variance of this population. So, this ratio multiplied by the $n - 1$ so this is also this total quantity is also one random variable, which is also a chi square distribution with parameter ν equals to $n - 1$ with this degrees of freedom $n - 1$.

So, this one is basically these two properties why I am stressing is that these two are the basic properties which is generally used for the different problems. The first one is that when we just take that difference from a particular value take the difference square

them and add them up. So, this property is generally used to test the goodness of fits. So, when we take a sample data and we take that how, what is the distance? from that, that target distribution. Suppose that I have the sample distribution and I can calculate their empirical cumulative distribution, and for some known standard probability distributions say for example, gamma distribution for that gamma distribution with some set of parameter I know how the CDF looks like.

So, what I can do with the from the sample data the calculated CDF and for a standard distribution there is in this case what I am discussing is gamma. So, that if we take the gamma distribution then what we can do for each point of the sample I can calculate what is the distance? of the two cumulative distribution function. And that distance, if I just square them or add them up and; obviously, then we can see from this discussion that, that particular random variable will follow a chi square distribution.

So, generally for the goodness of fit test that is why we use the chi square test. Secondly the another application is that, is that for this a square. So, when this is a square; that means, the sample estimate of the variance so for that, that is also follow a chi square distribution. So, if we need to know some properties of the sample estimate of the variance then also you can follow this distribution to have some statistical test. This will be discussed in more detail for this module seven while we are using make use of this chi square distribution and of course, the other distribution as well which we which I am going to discuss now.

(Refer Slide Time: 39:52)

Use of Chi-square Distribution in Civil Engineering

- Goodness-of-fit test for a sample data
- Statistical test sample variance

Probability Methods in Civil Engineering: H318 Dr. Rajib Maity, IIT Kharagpur 20

So, first of all I want to do this one the use of the title of the slide that use of chi square distribution in civil engineering. So, so far mostly for this all this distribution we have use this word that use of that particular distribution in civil engineering. But here I just draw your attention because that this use of this chi square distribution is need not be for the civil engineering only this is more general and this can be used for any particular application. Because this applications are more generally, more general in nature and this is the civil engineering is also not violated. So, even though I mention it here that in civil engineering, but this is in general for all the problems in hand. So, this two are we are just now we discuss goodness-of-fit test for a sample data that did not be only the civil engineering problem data that can be any data that is why, I just mention that one that is the general concept of this goodness-of-fit test and again the statistical test for the sample variance which also we discuss in this last slide.


(Refer Slide Time: 41:11)

Student's t Distribution

□ If Y is a RV following standard normal distribution and U is a RV following chi-square distribution with ν degrees of freedom, then

$$X = \frac{Y\sqrt{\nu}}{\sqrt{U}}$$

follows a t-distribution with ν degrees of freedom.

 Probability Methods in
Civil Engineering: M3L8

Dr. Rajib Halty, IIT Kharagpur

21

The next another important distribution is the student t distribution. If Y is a random variable following the standard normal distribution and U is a random variable following the chi square distribution with ν degrees of freedom. Then if we get another random variable like this that is Y by the square root of the chi square distribution normalized by its degrees of freedom, that is U by ν whole square root then, we will get this quantity which is another random variable and this is as then found to follow a student's t distribution this is known as t distribution or student t distribution with the degrees of freedom equals to ν .

(Refer Slide Time: 42:02)

Student's t Distribution...contd.


□ The pdf of t-distribution is given by

$$f_T(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu} \left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}} \quad -\infty < x < \infty$$

Where ν is a positive integer

□ The cumulative distribution is given by

$$F_T(x) = \int_{-\infty}^x \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu} \left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}} dx$$



Reliability Methods in
Civil Engineering: M3L8

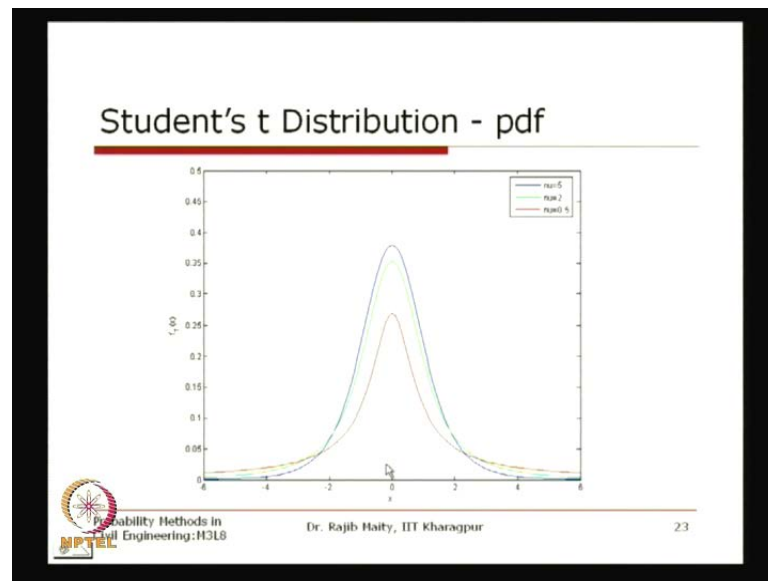
Dr. Rajib Maity, IIT Kharagpur

22

So, its distribution form looks like this the p d f of this t distribution is given by this form that is gamma of nu plus 1 divided by 2 divided by gamma nu by 2, this is the normalizing constant to make that mixture the total area under this p d f is equals to 1 multiplied by 1 by square root pi nu multiplied by 1 plus x square by nu power nu plus 1 by 2 and its support here is minus infinity to plus infinity which is here that. So, there is no bound of this support this is a entire real axis, over the entire real axis this random variable can take. So, this nu again is the positive integer which is the degree of freedom for this distribution. And to get this cumulative distribution again it is known that from this left hand support that is minus infinity in this case to specific value x this integration will give you the cumulative distribution of this p d f.

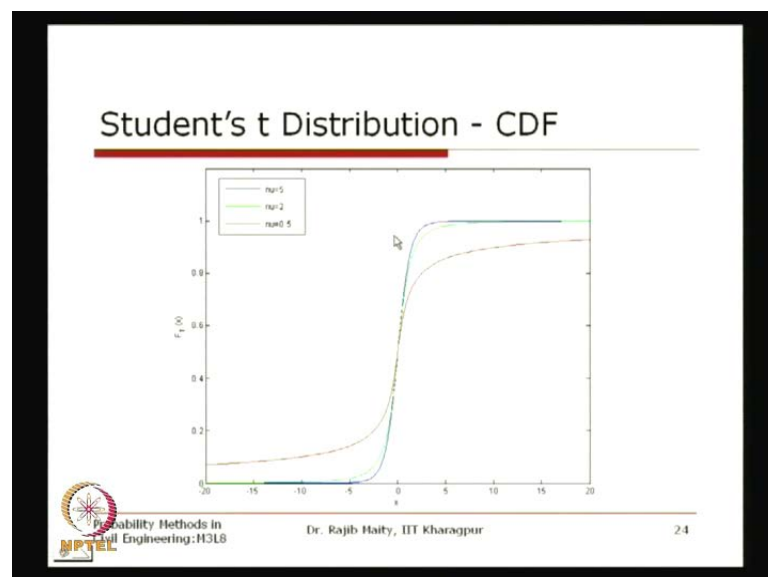
These are generally done numerically, because of close form of this integration is not available.

(Refer Slide Time: 43:21)



Now, again here we have shown different examples of this different values of this ν and you can see that here what we can see is that, if we increase the value of this ν this increase in peak is increases with the increase of this degrees of freedom and it can be shown that for the high value of this ν , high value of this degrees of freedom. It is becomes exactly similar to the normal distribution again. So, it will go and approach to this normal distribution with the increase of degrees of freedom that is what is shown here.

(Refer Slide Time: 44:04)



And for, the for those combinations of this degrees of freedom here we have shown that their distribution of this CDF that cumulative distribution of the student t distribution for different combination of that degrees of freedom. Here the red is the lowest then green and then highest is the blue with. So, far as this degrees of freedom is concerned.


(Refer Slide Time: 44:37)

Mean and Variance of Student's t Distribution

- Mean is given by

$$E(T) = 0$$
- Variance is given by

$$Var(T) = \begin{cases} \frac{v}{v-2} & \text{for } v > 2 \\ \text{undefined} & \text{otherwise} \end{cases}$$
- Skewness is zero (symmetrical) for $v > 3$



Reliability Methods in
 Civil Engineering: M3L8

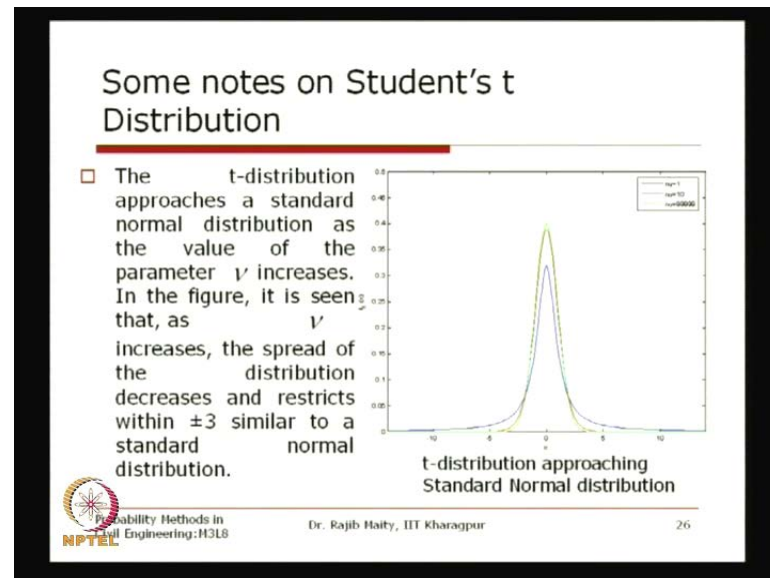
Dr. Rajib Haity, IIT Kharagpur

25

Now these are again the first few moments of this distribution the first of all the mean the mean of the students T distribution is equals to 0 when we are talking about this standard T distribution because this is the standard T distribution when we say. So, that is basically is being used when we do this statistical test and for this standard student t distribution, this you know that mean is 0 similar to the fact that for the standard normal distribution this mean is equals to 0 for this variance of this distribution is equals to is related to its degrees of freedom. That is nu by nu minus 2 and this is for when this nu is greater than 2 the degrees of freedom greater than 2 otherwise it is undefined. So you can see that this is a this is again. So, this as this new increases. So, this is also become this is this is, this variance is when this v is greater than 0 then you can say that we are having a positive variance; obviously, the variance is always positive so that is why when we're taking this for v less than 2 you know that. So, that is why this variance for the student t distribution for nu less than 2 is undefined and skewness is zero.

We know that we have seen that nature of this p d f. So, this skewness is zero as this distribution is symmetric. The skewness coefficient if you just see then this is generally defined for this ν greater than three, that is the degree of freedom greater than 3 we can say that this distribution is the skewness is zero.

(Refer Slide Time: 46:36)

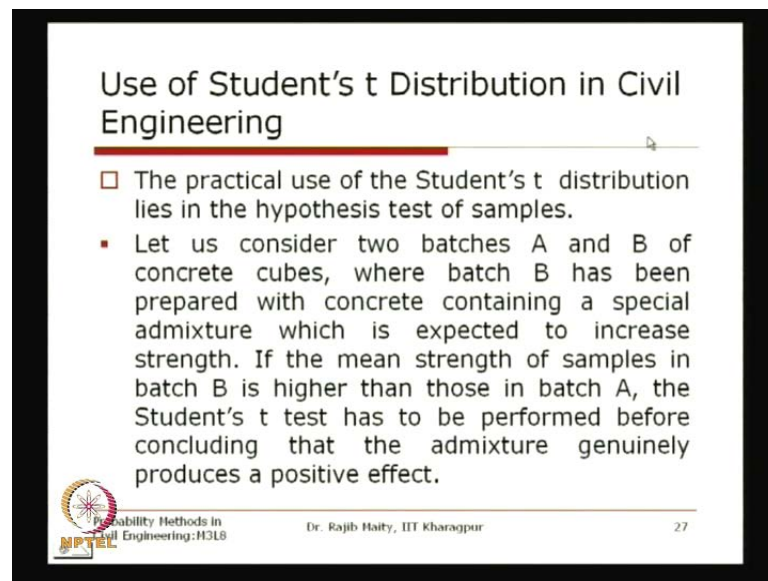


So, here if we see that this t distribution, this t distribution approaches to a standard normal distribution as the value of the parameter ν increases. In the figure it is seen that as ν increases the spread of the distribution decreases and restricts within this plus minus three to a similar to a standard normal distribution. This is what we are just discussing while giving this example of the effect of this different values of this ν . Here also there are three plots are shown: the first is blue one is for the lowest ν where in this case this ν equals to 1. In this case for this blue line and then, and then if we increase that ν that is for this red one you can see that it is ν equals to 10 for this green it is ν is a very high value is taken and shown that. So, it will mostly. So, it is not that it will go on in increasing it is, it is asymptotically with respect to this degrees of freedom it is asymptotically approaching to a standard normal distribution.

This is basically used in the case of this statistical test when you know that any sample data that is having some finite values of sample data. Now when we generally see that this dataset is the number of data points are less than, hence we even though we know


that particular statistics follow a normal distribution we generally go for this t distribution because for this t distribution that degrees of freedom is defined here and, we know that for the small number of this sample size instead of using that standard normal distribution we use the student t distribution which is more logical in this case because, for this standard normal distribution effectively the sample size should be infinity.

(Refer Slide Time: 48:46)



Use of Student's t Distribution in Civil Engineering

- The practical use of the Student's t distribution lies in the hypothesis test of samples.
- Let us consider two batches A and B of concrete cubes, where batch B has been prepared with concrete containing a special admixture which is expected to increase strength. If the mean strength of samples in batch B is higher than those in batch A, the Student's t test has to be performed before concluding that the admixture genuinely produces a positive effect.

 Reliability Methods in Civil Engineering: M3L8 Dr. Rajib Haity, IIT Kharagpur 27

So, use of student t distribution again I am discussing with respect to the civil engineering, but this applications application of all this three distributions student t distribution, chi square and F distribution are more general in nature. The practical use of student t distribution lies in the hypothesis test of this sample particularly, for the sample mean. So, two samples are available to us and if we have the two different mean then then the target is the whether, we have to estimate that whether the which mean is more, which mean is less. So, if that kind of thing that I cannot simply calculate the sample mean and say that the more and less by, we have to test this statistically in the sense that, that whether that with the some statistical significance level whether the one mean is really larger than others are not.

So, one example from the civil engineering application we can take that take in the in case of this admixture in the concrete sometimes we mix some admixtures to a concrete mix and claim the strength as increased. So, in that kind of test what we have to do is that we

have to do the two different batches of this concrete cube in one batch that admixture will be added and other batch without that particular admixtures. Then we have to take that sample of the calculate that sample mean and then, if you want to test that which mean is more whether that is that, that is an effect of this mixing this admixture are not then that, we have to test that hypothesis we have to test while testing this kind of hypothesis where the mean is involved then we have to follow the student t distribution.

This is what is explained here that let us consider two batches A and B of the concrete cube where the batch B has been prepared with the concrete containing a special admixture which is expected to increase its strength. If the mean strength of the sample in the batch B is higher than those in the batch A the student t test has to be performed before concluding that the admixture generally produce a positive effect or not. So, it is not only for this one. So, wherever we have we can say that there is a that mean is in further sample mean is involved we use this student t distribution.


(Refer Slide Time: 51:25)

F Distribution

□ If U and V are independent chi-square distributed RVs with m and n degrees of freedom respectively, then the RV

$$X = \frac{U/m}{V/n}$$

is a F-distributed RV having degrees of freedom $\gamma_1 = m$ and $\gamma_2 = n$.
 Here F is sometimes called the variance ratio and $\gamma_1 = m$ and $\gamma_2 = n$ are called the numerator and denominator degrees of freedom respectively.



Reliability Methods in
 Civil Engineering: M3L8

Dr. Rajib Halty, IIT Kharagpur

28

The last one in this series is the F distribution if the u and v are independent chi squared distributed with m and n degrees of freedom. Then their ratio that is U by m and v by n we will get a new random variable which is F distributed and for the F distribution they are two degrees of freedom and here. So, that is γ_1 and γ_2

as denoted is here the gamma 1 equals to m and gamma 2 equals to n. Here the f is sometime called as variance ratio and this gamma 1 equals to m and gamma 2 equals to n are called the numerator and denominator degrees of freedom respectively.

(Refer Slide Time: 52:05)

F Distribution...contd.

□ The pdf of F Distribution is given by

$$f_F(x) = \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right) \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{\Gamma\left(\frac{v_1}{2}\right) \Gamma\left(\frac{v_2}{2}\right) \left[1 + \left(\frac{v_1}{v_2}\right)x\right]^{\frac{v_1 + v_2}{2}}} \quad x \geq 0$$


= 0 otherwise

Where v_1, v_2 are positive integers

□ The cumulative distribution function is given by

$$F_F(x) = \int_0^x \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right) \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{\Gamma\left(\frac{v_1}{2}\right) \Gamma\left(\frac{v_2}{2}\right) \left[1 + \left(\frac{v_1}{v_2}\right)t\right]^{\frac{v_1 + v_2}{2}}} dt \quad x \geq 0$$

= 0 otherwise



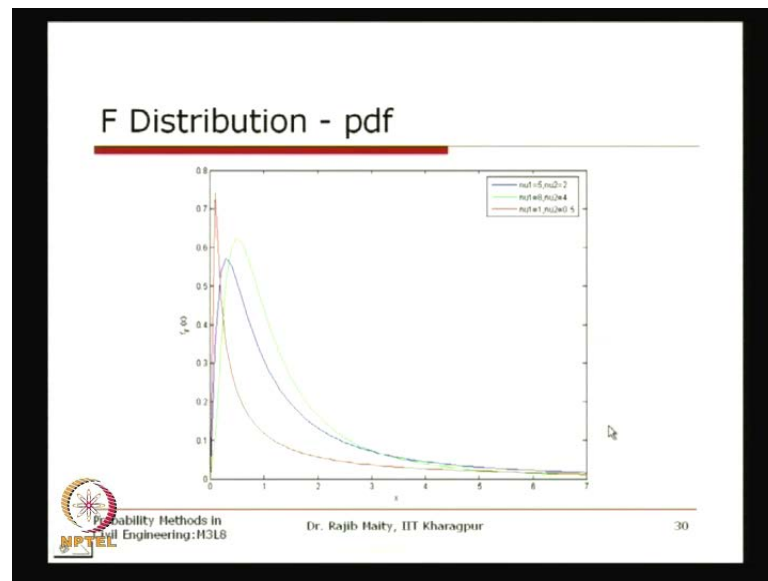
Reliability Methods in
Civil Engineering: M3L8

Dr. Rajib Maity, IIT Kharagpur

29

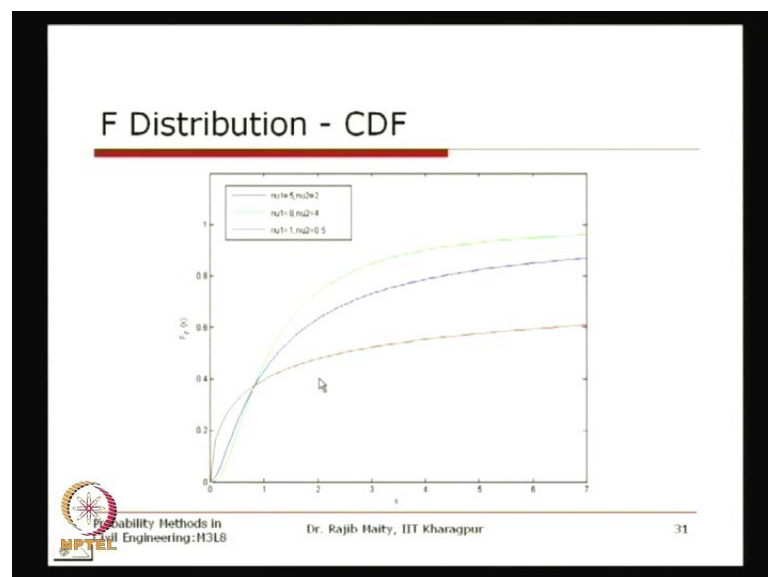
So, this is the general form of the pdf of the F distribution that is gamma nu 1 plus nu 2 by 2 divided by gamma nu 1 by 2 gamma nu 2 by 2 nu 1 by nu 2 power nu 1 by 2 this is the full the normalizing constant, multiplied by x power nu 1 minus 2 by 2 1 plus nu 1 by nu 2 into x power nu 1 plus nu 2 by 2 and this is for this x greater than equals to 0 and otherwise it is 0. And its cumulative distribution again at the close form is not available to go for some numerical integration.

(Refer Slide Time: 52:48)



Now, this is the effect of these different values of this combination of two degrees of freedom. One is the ν_1 and ν_2 . We can see that thus red, blue and green different combinations are shown here for different set of combination of degrees of freedom.

(Refer Slide Time: 53:09)




And their respective cumulative distribution functions are shown here which is starting; obviously, from 0 and it is asymptotic to 1 starting from 0 going and asymptotic to 1.

(Refer Slide Time: 53:37)

Mean and Variance of F Distribution

- Mean is given by
$$E(F) = \frac{\gamma_1}{\gamma_2 - 2} \quad \text{for } \gamma_2 > 2$$
- Variance is given by
$$\text{Var}(F) = \frac{\gamma_1^2(\gamma_1 + 2)}{\gamma_1(\gamma_2 - 2)(\gamma_2 - 4)} \quad \text{for } \gamma_2 > 4$$


 Probability Methods in
Civil Engineering: M3L8 Dr. Rajib Haity, IIT Kharagpur 32

The mean of this F distribution is $\frac{\gamma_1}{\gamma_2 - 2}$ for $\gamma_2 > 2$ and this variance is defined by $\frac{\gamma_1^2(\gamma_1 + 2)}{\gamma_1(\gamma_2 - 2)(\gamma_2 - 4)}$ for $\gamma_2 > 4$.

(Refer Slide Time: 53:50)

Use of F-Distribution in Civil Engineering

- The practical use of the F-distribution lies in the statistical test of variances of two samples.
- Let us consider that n_1 and n_2 number of tests are performed by two separate methods for the analysis of a constituent in an effluent. If the two methods yield different values of variances, the F-distribution may be used to decide which method is more precise at the 1% and 5% levels of significance.

 Probability Methods in
Civil Engineering: M3L8 Dr. Rajib Haity, IIT Kharagpur 33

This use of F distribution here are in civil engineering and mostly it is for thus general case when now for this t distribution we told that when the mean of the sample is

involved that time we use the t distribution similarly and when the sample variance is involved we here use this chi-square distribution now suppose that there are two samples are available and two the ratio of the sample variance from the two samples are concerned then that ratio to test that ratio we have to follow that F distribution. Suppose that one sample is having one standard deviation is there which is estimated from the sample data. Similarly, for the other one now, if I take ratio and if I want to test which 1 is more than; that means, that ratio should be either greater than 1 or less than 1 are equal to 1 say if equal to 1 then both the samples, both the samples' standard deviation is same and if it is more depending on which one is more and less that can be decided.


So, for the ratio of the variances from two population follow a F distribution. Keeping that in mind we can just see where this will be effective the practical use of F distribution lies in the statistical test of the variance of two samples. Let us consider that n_1 and n_2 number of test are performed by two separate methods of this analysis of a constituent in an in an effluent if the two methods yield different values of this variance the F distribution may be used to decide which method is more precise at 1 percent and 5 percent level of significance.

So, this level of significance and which one is standard we will be discuss in more detail in the subsequent module and that last module seven when we discuss about the statistical test.

(Refer Slide Time: 55:45)

Use of F-Distribution in Civil Engineering...contd.

□ Let us consider that the distribution of rainfall depths received by two catchments. If the mean depth is equal for the two catchments but the variances differ, then the F test can be used to test the ratio of the variances and conclude whether the variation of rainfall depth are statistically equal. If not, then which one is having more variation.



Probability Methods in
Civil Engineering: H318

Dr. Rajib Haity, IIT Kharagpur

34

And in civil engineering specially, we can also say that let us consider that the distribution of the rainfall depth received by two catchments are there, if the mean depth is equal for this two catchment, but the variances differs, then the f test can be used to test the ratio of the variances and conclude whether the variation of the rainfall depth is statistically equal are not. If not then which one is having the more variation.

So, with this I mention once again that this may not be the complete list of this all the standard probability distributions, but these are the mostly used into address the different problems in civil engineering. Thank you.