INDIAN INSTITUTE
OF
TECHNOLOGY
KHARAGPUR

NPTEL
National Programme
on
Technology Enhanced Learning

Probability Methods in Civil Engineering

Prof. Rajib Maity

Department of Civil Engineering
IIT Kharagpur

Lecture – 13

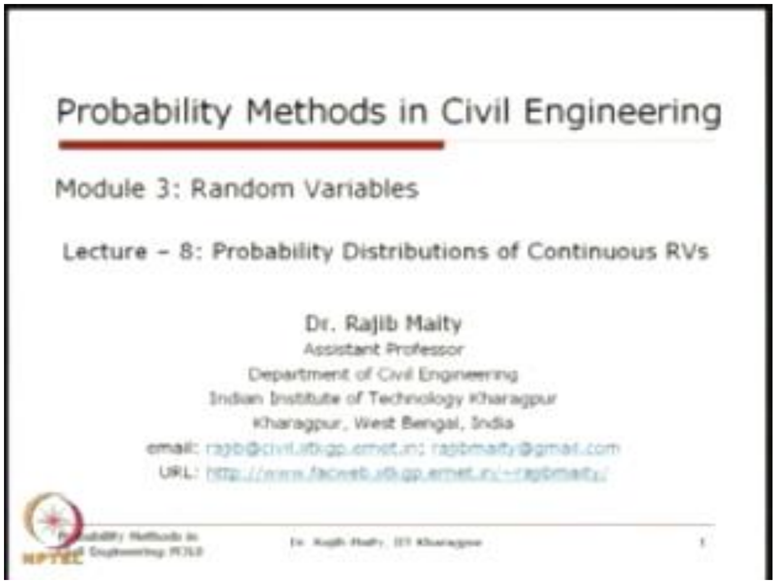
Topic

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.

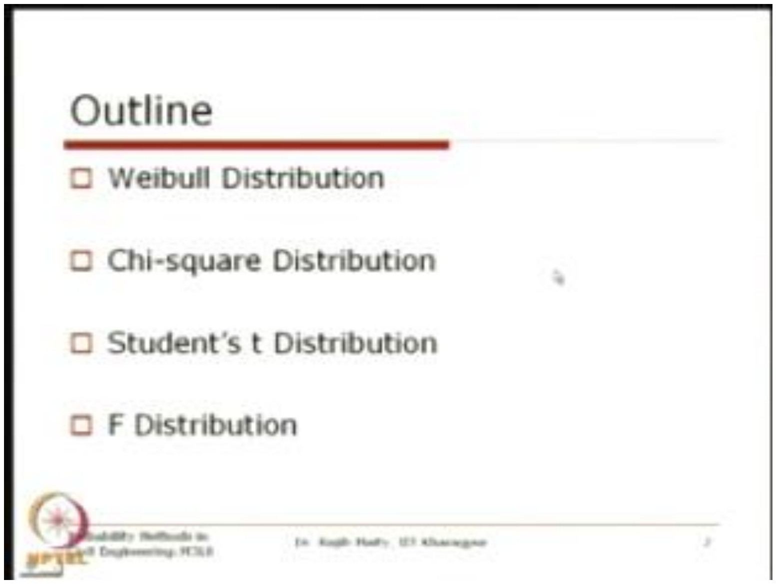
But these are the distribution which are mostly used for different problems in civil engineering. So, in the last class we, there is the Weibull distribution was pending. So, we will start with the distribution this class, and we will also discuss some of the distribution after that we mentioned in the last class that chi-square distribution, t-distribution, and f-distribution which are mainly used for the different statistical test.

(Refer Slide Time: 01:30)



So, our this lecture on that probability distributions of continuous random variable we will continue in this lecture as well.

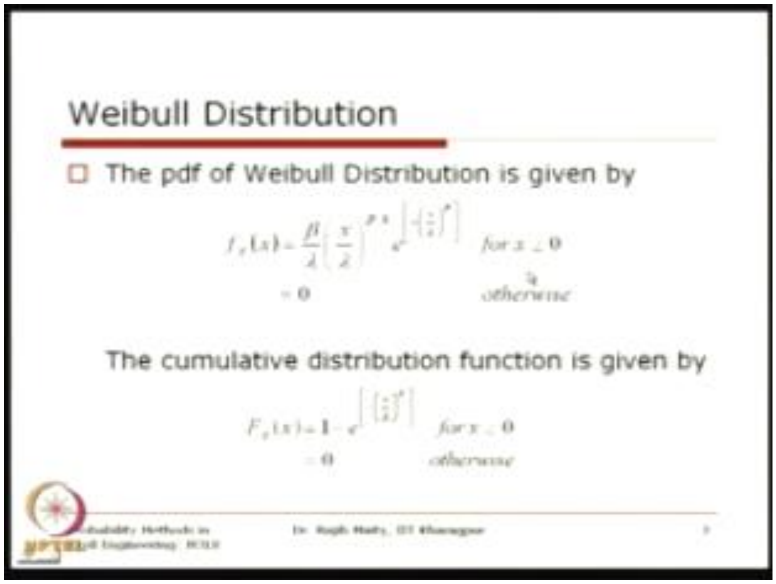
(Refer Slide Time: 01:39)



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.

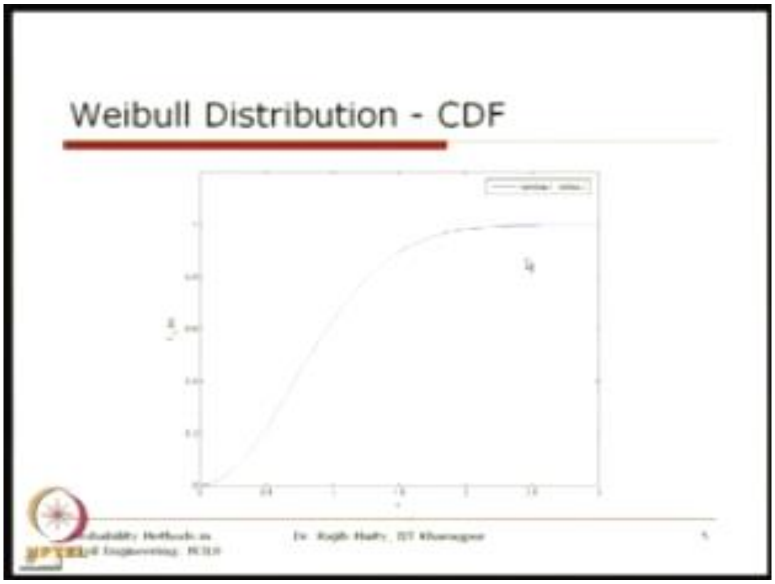
(Refer Slide Time: 02:29)



So, we will start with this weibull distribution this general form of the pdf of weibull distribution looks like this there is two parameters one is β and another one is λ. So, this pdf probability density function is β/λ $x/\lambda^{\beta-1}$ exponential minus x/λ^{β} for x greater than equal to 0. So, for this support of this distribution is bounded at 0, so, lower bound is 0and upper bound is infinity. And for the negative values of x this is not defined this is 0 and we know that thus cumulative distribution for any pdf is the integration from this lower support that is here is 0 up to that particular value x.

If we do that integration we get the cumulative distribution function and this is this can be shown that, this is $1-e^{-x}/\lambda^{\beta}$ for x greater than equal to 0. So, if we compare this distribution from the whatever the distribution we have discuss earlier then we will see that we discuss earlier the exponential distribution, we discuss earlier that γ 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 γ distribution.

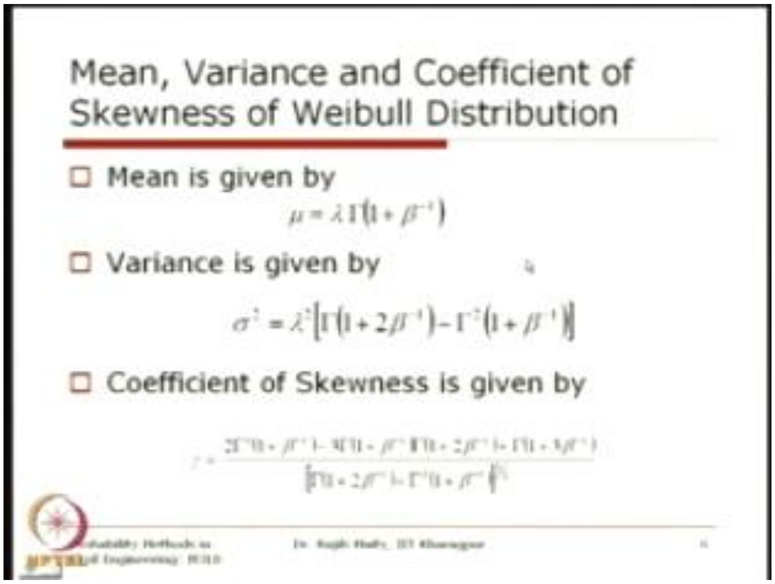
(Refer Slide Time: 04:18)



So, the other properties of this distribution is means first this is one example of that weibull distribution pdf, how the pdf look like here the parameters that is taken as this λ=1 and β=2 so, taking this to, this set of parameter this is the general shape of this one which we can see that this 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. And for the same combination of the parameters that is λ=1 and β=2 that is same pdf that is shown in last slide if we take that cumulative distribution that is 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: 05:26)

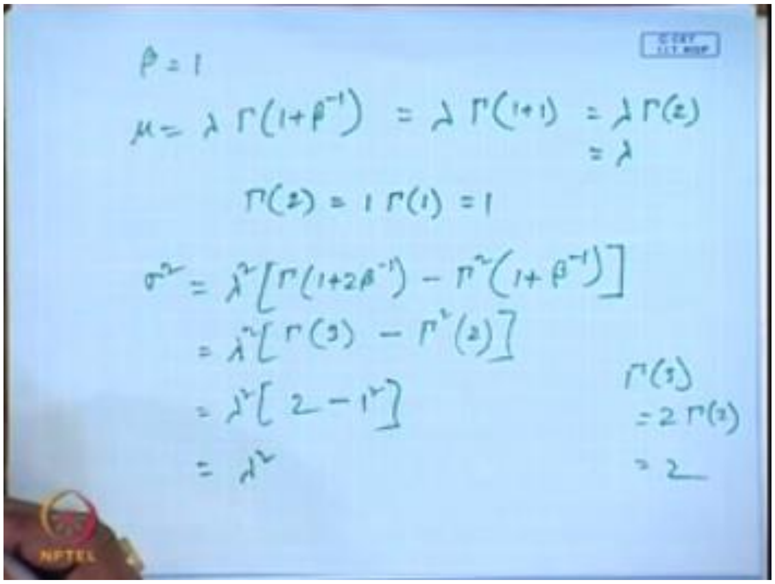


So, now this first few moments if we discuss about this one that this mean is given by $\mu=\lambda$ this is one parameter and this γ function of $1+\beta^{-1}$ that is $1+1/\beta$.We discuss about this γ function some previous lectures while discussing that beta distribution and gamma distribution. 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 λ^2 x $\gamma(1+2\beta^{-1} – \gamma^2 1+\beta^{-1})$ basically this one if we just this λ^2 if we take inside this bracket this becomes that mean. So that mean square so we can say that this quantity that is the λ^2 x $\gamma 1+2\beta^{-1}-\mu^2$ will be the variance of the distribution. And similarly, if we take that skewness we will gets along expressions likes 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 $\beta=1$ then, you know that this will become that $\gamma 1+\beta^{-1}$.

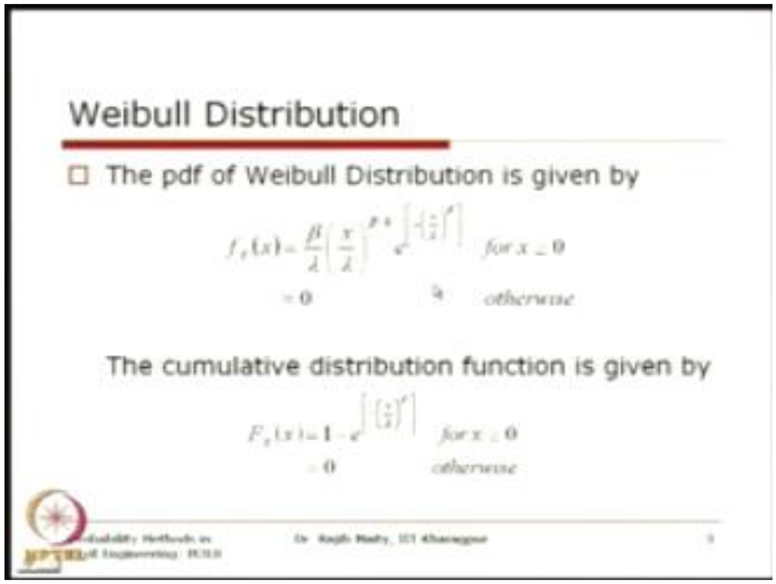
(Refer Slide Time: 07:14)



So, what we are trying to say here if we just put that $\beta=1$, then that μ that λ, $\gamma 1+\beta^{-1} = \lambda\gamma 1+1 = \lambda\gamma 2$. So, and from the earlier lectures we know this gamma if this1is 1 integer then, we know that this if $\gamma n=n-1\ \gamma n-1$ 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 μ becomes λ. Similarly, if I take that sigma square which is the expression is this λ^2 x $\gamma 1+2\beta^{-1}-\gamma^2 1+\beta^{-1}$.

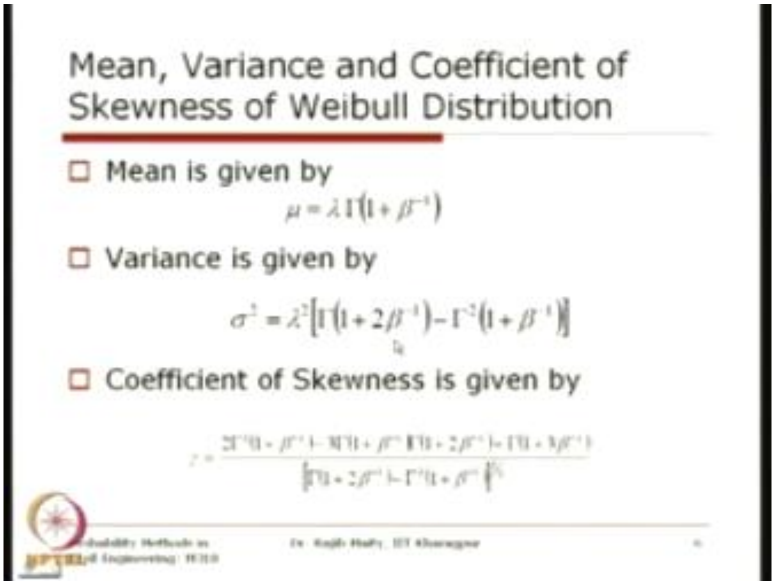
So putting this $\beta = 1$ it will eventually come that λ this $\lambda^2\gamma$ this is 3- γ^2 (2).So, this is 1 and this will become your 2.So, λ^2 this is 2. So, $\gamma 3$ you know $\gamma 3 = 2\ \gamma 2$ and $\gamma 2 = 1$. So it is $2 - 1^2$. So this is equals to λ^2 now if we see that for the exponential distribution we got exactly the same value of this moments at least the first moments that is for a exponential distribution that mean is equals to λ and variance is equals to λ^2 so, if we put this $\beta = 1$ it seems that it is a it is becoming an exponential distribution.

(Refer Slide Time: 09:13)



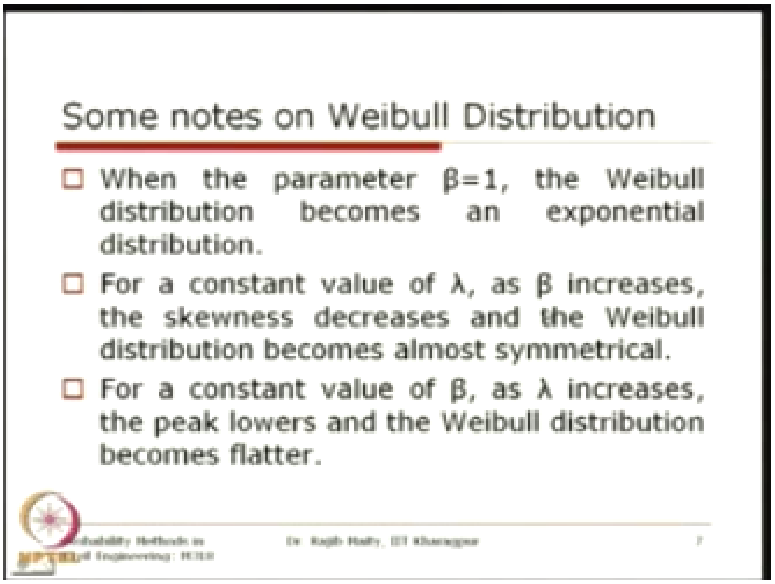
So, if we just go back to the to this distribution of this weibull pdf and here if you put that $\beta = 1$ then we can see that that is the $1 / \lambda$ and here it is $\beta = 1$. So, this is 1 and this e^{-x}/ λ for $x > 0$ and otherwise it is 0 so which is nothing but exactly same to this exponential distribution. So, here so with respect to the exponential distribution we can say that one more parameter is introduced which is β if this $\beta = 1$ this weibull distribution becomes exponential distribution and we will also show that how this weibull distribution is also that γ distribution is a special case for this weibull distribution that we will discuss in a minute.

(Refer Slide Time: 10:04)



So now this we discuss and this first few moments mean, variance and coefficient of skewness we discuss.

(Refer Slide Time: 10:13)

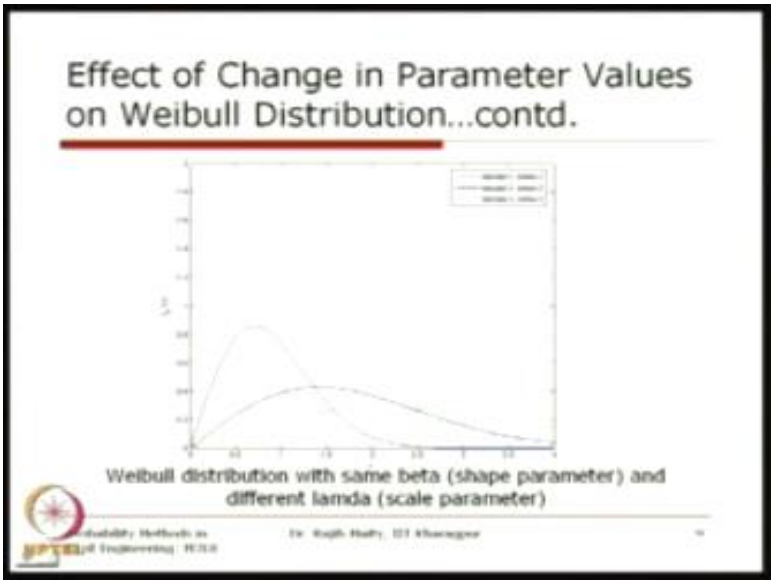


Some note that what just now, we discuss with the first note that when parameter $\beta = 1$ the weibull distribution becomes an exponential distribution which just now we discuss. Now if we just want to see the effect of the parameters that is both for this λ and β. So, if we just keep one parameter constant and if we just the increase the other parameter then what will be the effect on the shape of this pdf of weibull distribution. So the first it says that for a constant value of λ. So if you just take constant value of λ and if we increase this β.

So as β 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 symmetrical 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 β 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 β.

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

(Refer Slide Time: 12:23)



So this is the first thing when we are keeping this λ constant and adding this β and β = 1, 2 and 4 and keeping the all main all three cases λ =1.What you can see here that this for this blue line this pdf shown by blue line this is for the combination that λ= 1 and β = 1 for this black one the combination is the λ = 1 and β is = 2 and for this green one this λ = 1 means all the λ are same for all three and β = 4.So, from this blue then black, then green the β 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 β = 1when we say that β = 1 that means, we are talking that irrespective of any value of λ so when β = 1 it approaches to the exponential distribution. So that we can see here that this is an this an exponential distribution and we know that this at 0 at x = 0 the value of fx is1 / λ so as we are taking the λ = 1 that is why that x = 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 have discussed in previous lectures.

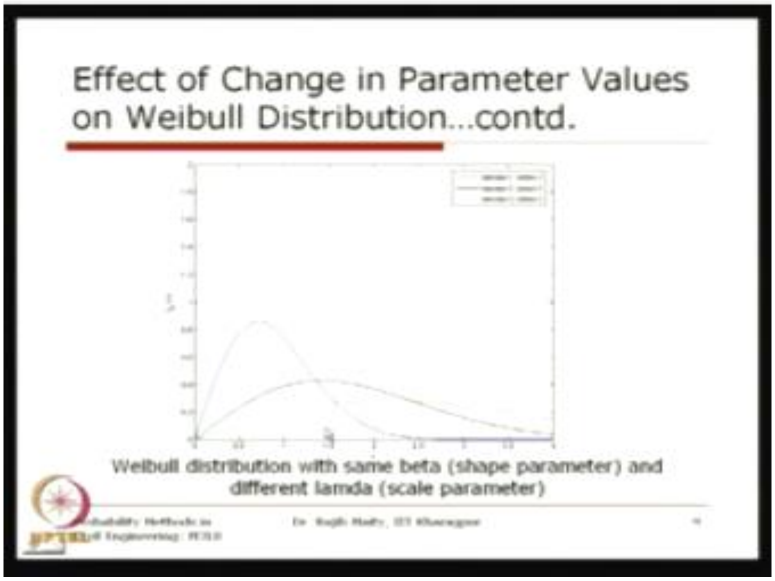
Now, when we increase β from $\beta = 1$ to 2 the skewness that asymmetry decreases. So it approaching to the symmetrical and it is more clear when we are going to this four that it is becoming some more to more symmetrical compared to the other two graphs where $\beta = 1$ and $\beta = 2$. So, if we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

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 pdf's with different combination of λ and β. So, this first one this blue one is for this $\lambda=1$, $\beta = 2$ this black one is the $\lambda = 2$, $\beta = 2$ and this green one is for $\lambda = 3$ and $\beta = 3$.Thus for all three pdf's here the β is same which is equals to 2 and for the and the λ is increasing from1 then for the black it is 2 and for green from it is 3.

So what we can see is that when this λ is low then it is peak is more and with the increase of this λ this peak becomes flatter. Basically, when we are having some set of data there are some test 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 parameters are 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.

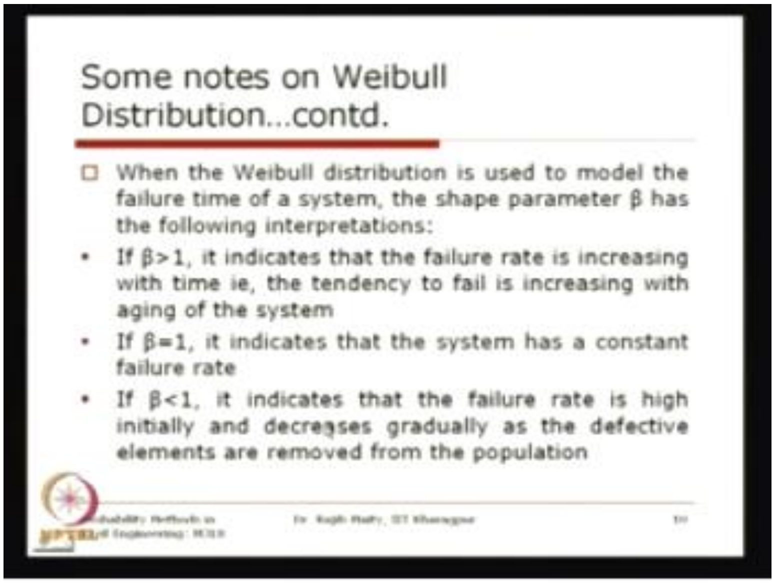
That which distribution this is following and then we estimate the parameter. So that while estimating the parameter there are different methods are available. So here what we are discussing that the effect of this parameter the graphical significance of this parameter if we select one parameter if we change one parameter how the shape and the location of particular pdf changes that is what we are discussing for all the models.

(Refer Slide Time: 16:57)



All the probability distributions.

(Refer Slide Time: 17:02)

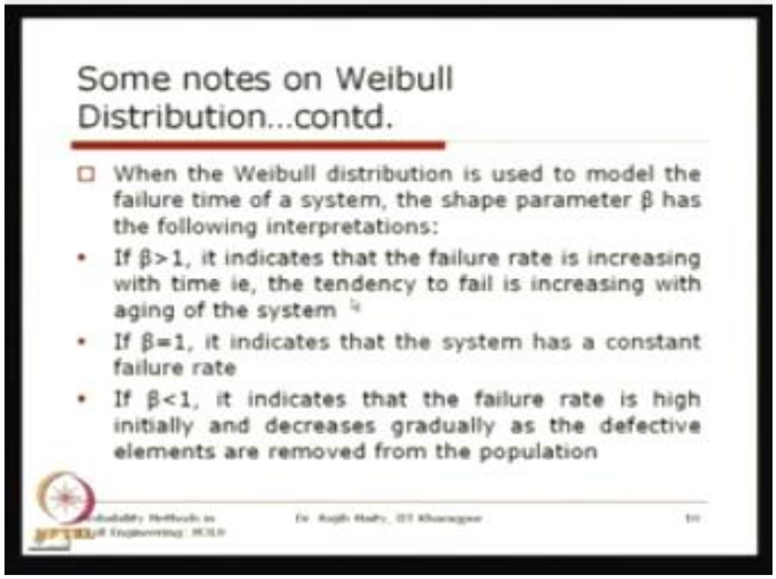


Now one important discussion with respect to this Weibull discussion is this the parameter beta, this parameter beta is playing a vital role and it is having different meaning for the different ranges of this parameter first. So when the Weibull distribution is used to model the failure time in a system 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 $\beta = 1$ that is the when we say that $\beta = 1$ 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 inter arrival rate average time on which a particular a 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 e $\beta = 1$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 $\beta = 1$ the distribution is approaching to the exponential distribution. Now there are two other region if $\beta< 1$ or if $\beta > 1$, what does it implies?.

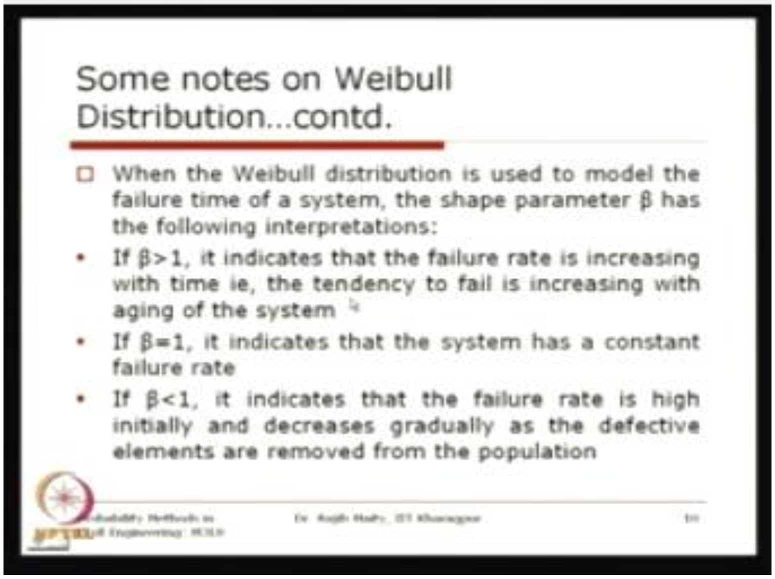
(Refer Slide Time: 18:53)



So when we say that this $\beta < 1$ 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 apparatus for its fatigue problem when the time increases sometimes the inter arrival rate of the phenomenon may increase.

So here we are deviating for 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.

(Refer Slide Time: 19:46)

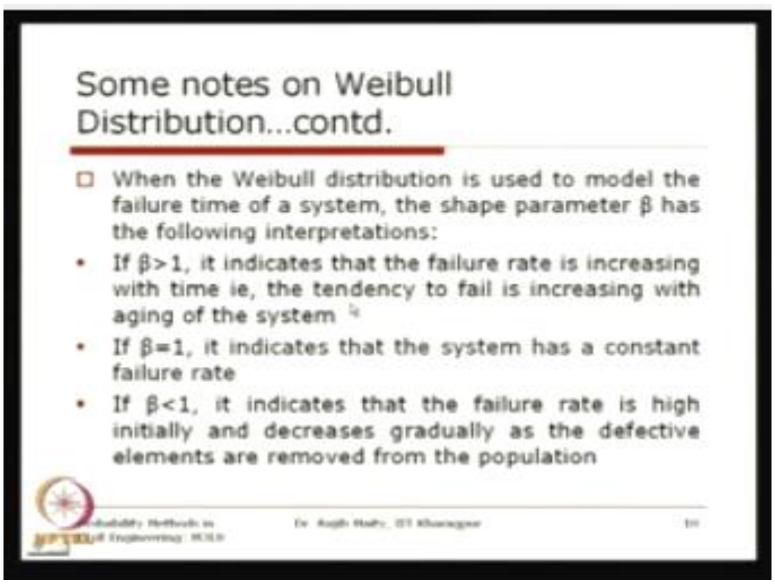


So when we say that this rate of failure is rate of failure is increasing with time that means that is captured when the distribution is having the parameter that $\beta > 1$. Similarly so this is you know that when some particular structural member or the structure itself due to its aging problem or when it is passing more time the chances of failure is increases, so that kind of cases this weibull distribution with $\beta > 1$will be the suitable case.

On the other hand when the $\beta < 1$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 $\beta \leq 1$.

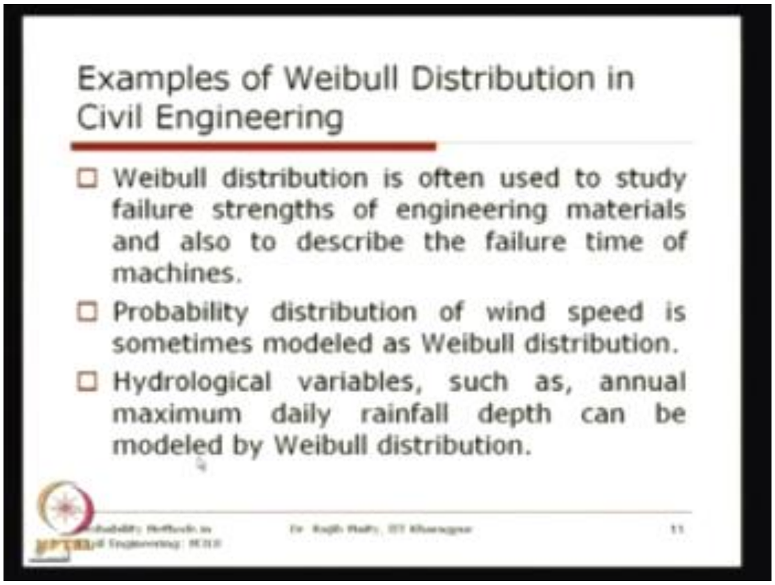
So at the birth is shown the mortality rate is more compared to when the age increases of this of the child. Similarly for if we come to this some application of civil engineering that time when we are producing some kind of product and we are just marketing it. Now if there is any defective element 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 to the initial time is more compared to that moment increases. So at that case that $\beta < 1$.

(Refer Slide Time: 22:13)



So that that the failure rate is high initially and decreases gradually as time pass. 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.

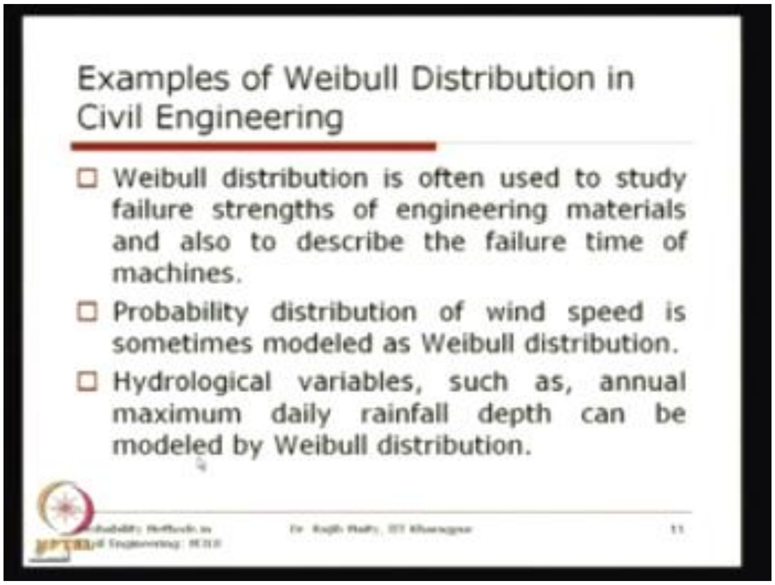
(Refer Slide Time: 22:35)



This 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.

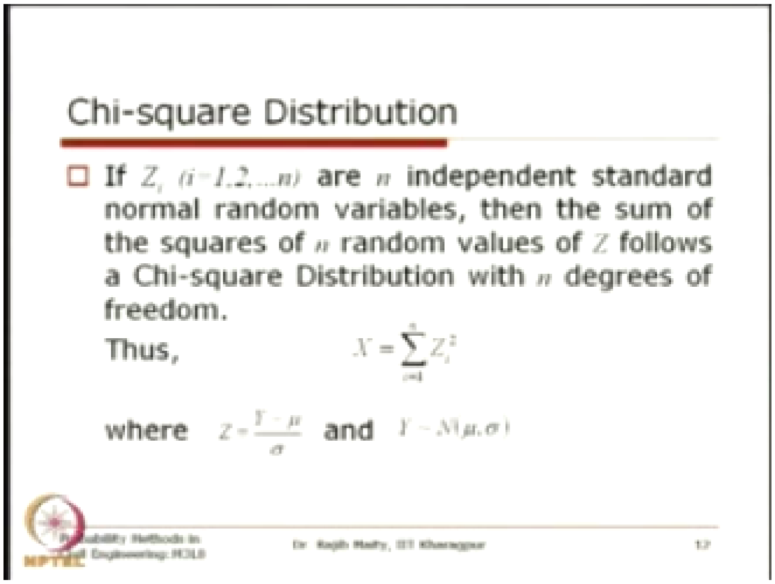
That means weibull distribution with parameter $\beta = 1$ or it depends on that what kind of situation, what kind of problem in hand based on that the beta parameters may change.

(Refer Slide Time: 23:19)



Secondly the sometimes the probability distribution of the wind speed particularly the hourly wind speed is also has 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 and that data also has found to follow a weibull distribution with suitable combination of its parameter λ and β.

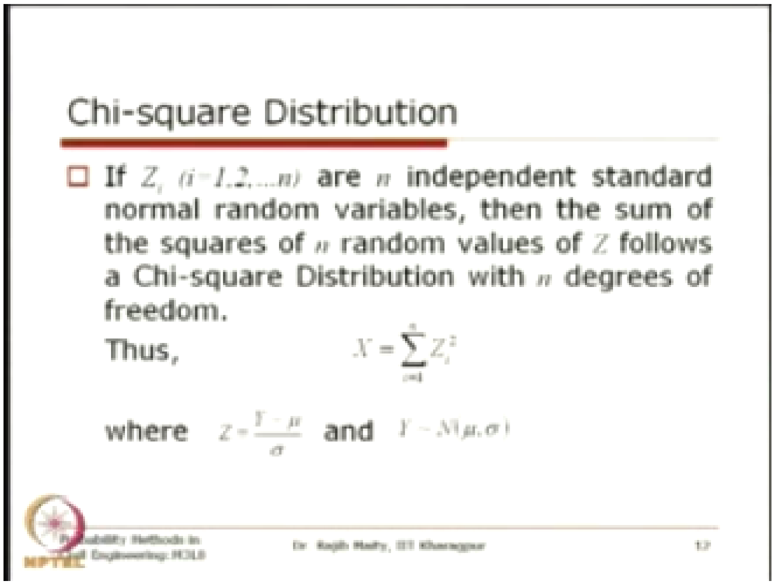
(Refer Slide Time: 23:57)



Next we will discuss three distribution one after another that we will start with the chi- square distribution and this 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 were discussing this л√ distribution which is generally used for the purpose of for the purpose of goodness of fit so many times I have mentioned that ones the sample data is available to us.

We generally have to first we have to test that distribution it is following and so they this is also can be tested statistically and that test is generally done through this л√ test it is called and based on this л√ distribution so we will understand now the what are this distribution and what are the different properties so first we are starting.

(Refer Slide Time: 24:55)

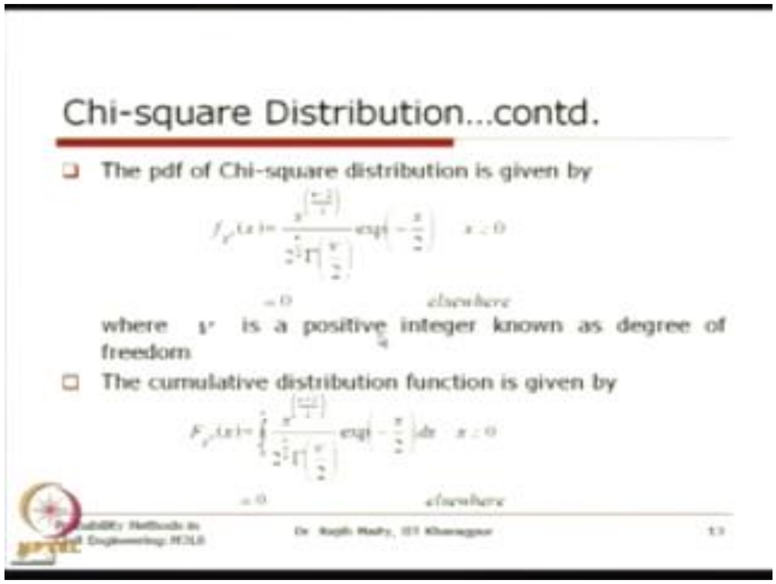


This Chi - square distribution if Z^i = 1 to n that is Z1 Z 2 Z3 up to z n are independent and identically distributed of course we use the word this 3, 3 were together independent and identically distributed sometime it is abbreviated as iid here these are independent of course and following the standard 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 degrees of freedom so here that x that the random variable that I am defining is equals to the summation of those $Z1^2$ + $Z2^2$ + up to Zn^2 so where this Z are the standard normal distribution so if I say that his Y is following a normal distribution with μ and α their parameter then we can make it we can get their reduced variate Z = Y – μ by α and this Z for this n numbers of independent random variables are there.

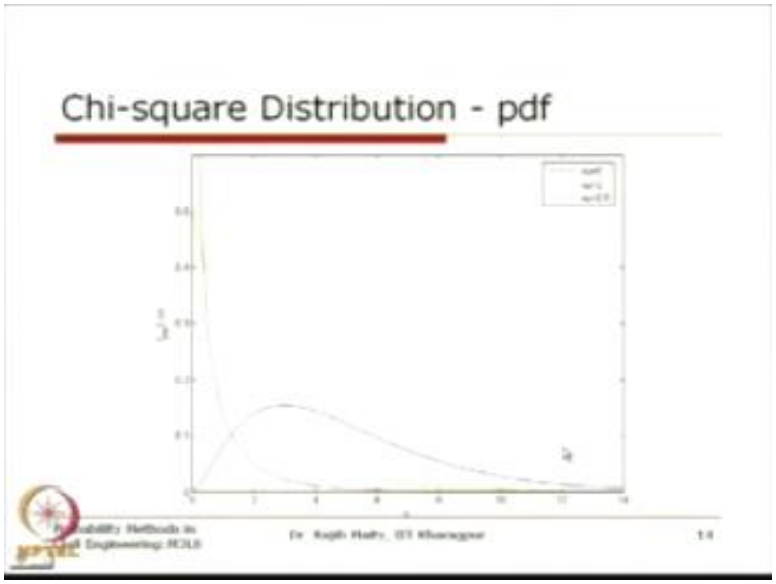
If we square them of add them then these random variable that summation x that is follow a chi square distribution now the pdf of this chi - square distribution.

(Refer Slide Time: 26:42)



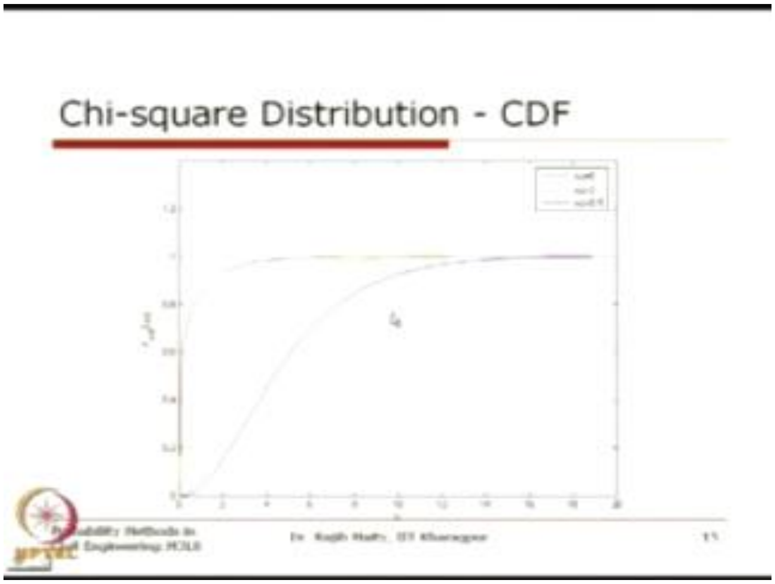
Is like this that is this is f chi square x is x power μ - 2 by 2divided by 2 power μ by 2 Γ μ by 2 exponential - x by 2for x greater than equals to 0 here μ is the parameter known as the degrees of freedom this is an integer it is an positive integer and this distribution is again lower bound by 0 and the for negative values it is 0 so this μ here mentioned that this μ 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 pdf and this is also defined for this x greater than equal to 0 and 0, 0 elsewhere so we will know that it's.

(Refer Slide Time: 27:48)



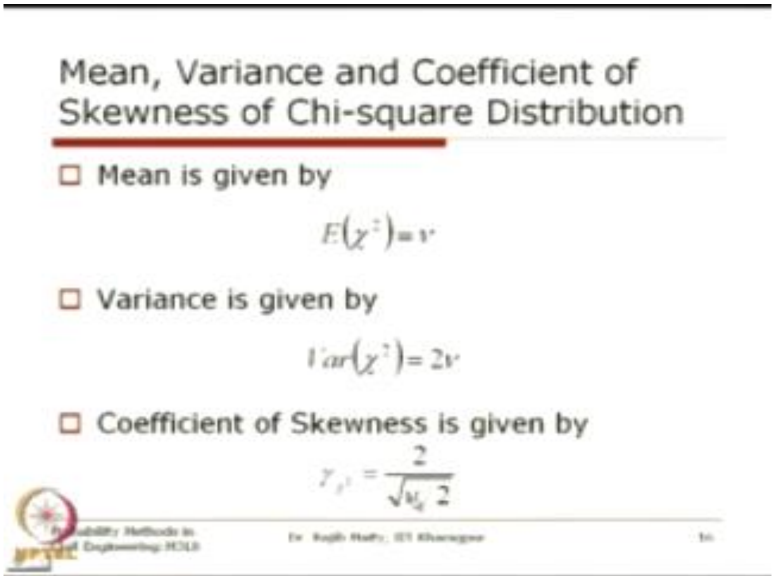
Different moments and different initial moments now before that we will show the how this for this different values of this μ which is that the degrees of freedom for this chi square distribution is this so here we have used different combination of this nu and we are so showing that as this μ increases then this is becoming this so this blue is for the maximum values of new 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: 28:38)



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.

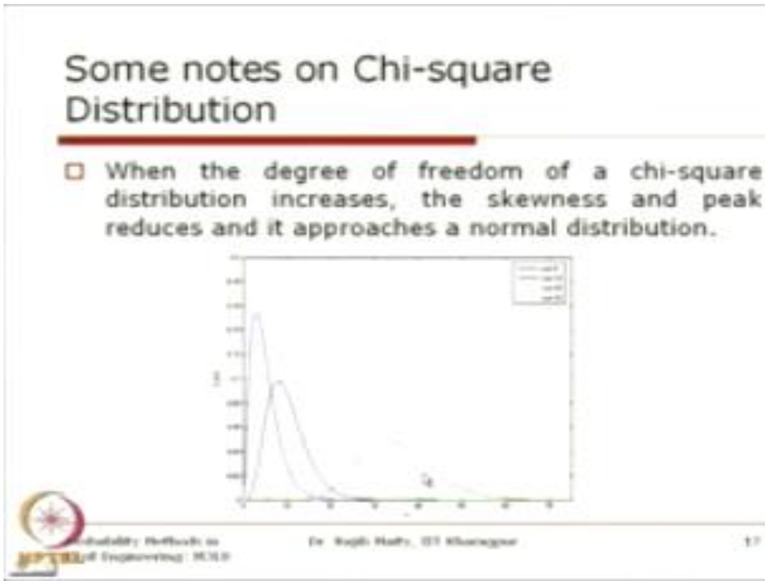
(Refer Slide Time: 28:56)



Now the initial few moments if we see then there are some interpretation particularly it is skewness as just we have seen from this graph the first is this its mean the mean of this chi square distribution is equals to its degree of freedom which is nu the variance of the chi-square distribution is 2 μ that is double of the mean and the coefficient of skewness if we see then this is that 2√ μ by 2 so this√ μ 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 μ that is if the degree of freedom is increased the μ 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 μ.

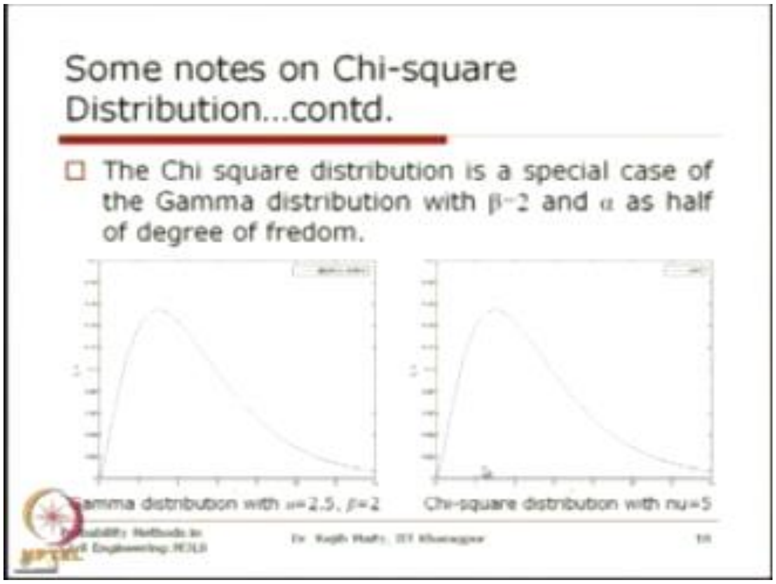
(Refer Slide Time: 30:12)



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 is approaches to a normal distribution in the sense that it is approaching to a symmetrical distribution and going away from it is 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 μ = 5 the black one that μ = 10 then this green one is 20 and this μ is 35.

So here you can see that as this degrees of freedom is increasing this distribution becomes almost symmetrical of course all this 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.

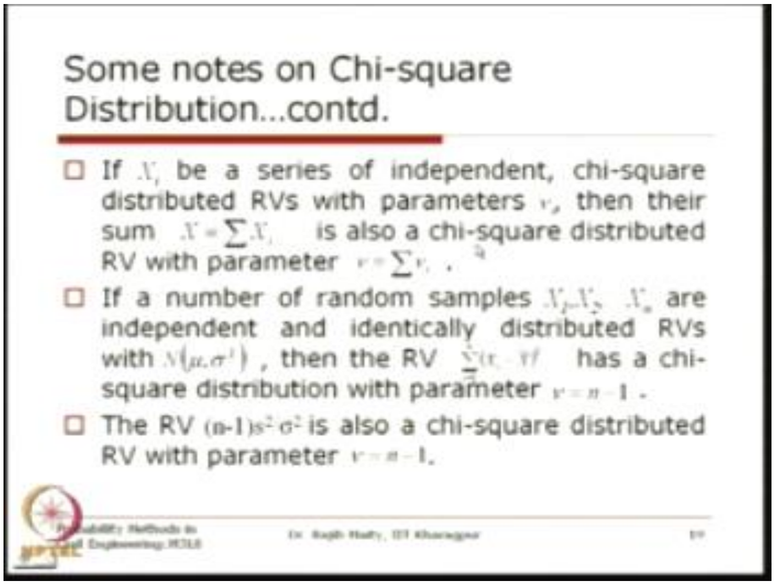
(Refer Slide Time: 31:31)



Now this chi square distribution is a special case of the Γ distribution with β = 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 if you see this is chi square distribution with μ that that is the nu the parameter of this chi square distribution is equals to 5. So, if we take this one that nu=5then the shape look like this the pdf 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 these degrees of freedom that is 2.5.

So, if we plot a gamma distribution with α=2.5 and β=2, then we get the exactly the same pdf 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.

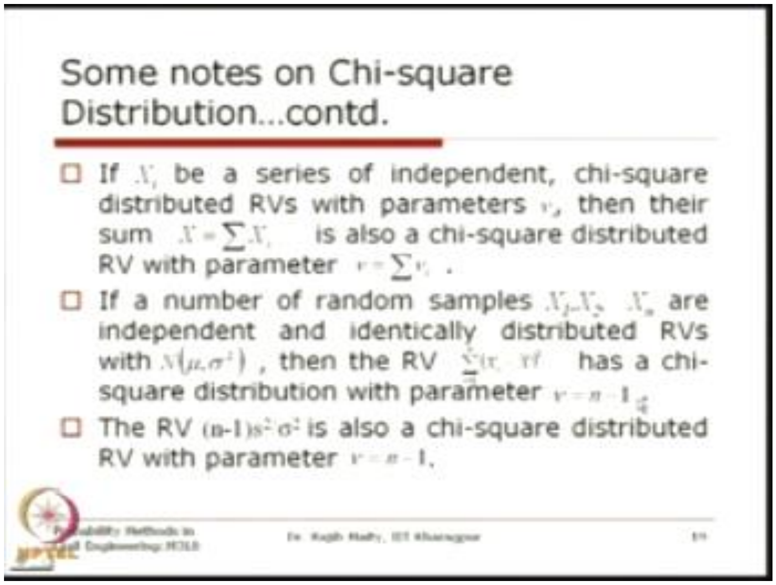
(Refer Slide Time: 32:50)



Now, if that X_i be a series of independent chi-square distribution, chi square distributed random variables with parameter nu_i that is for first that X_1 this degrees of freedom that the parameter is nu1 X_2 parameter is nu 2 then their some, if I take the some of those all those X_i is that is from 1 to n, then the a new random we will get a new random variable X which is the ∑ of them, simple arithmetic summation.

Then this new random variable is also a chi square distribution with the parameter is equals to ∑ 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 ∑ of the degrees of freedom for all those random of all those random variables which we added up. So, that is explained here for this first note of this chi square distribution.

(Refer Slide Time: 34:13)



The second is the, if a number of random variable that is X_1,X_2, 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 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 nu=n-1.

Now, there are two things here the first one is that we have in random variable 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 equals to the ∑ of all this 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 one 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 nu=n, because we are adding n degrees of freedom, but here the degrees of freedom is that n-1. So, this 1degrees 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 they from them to and then square it up.

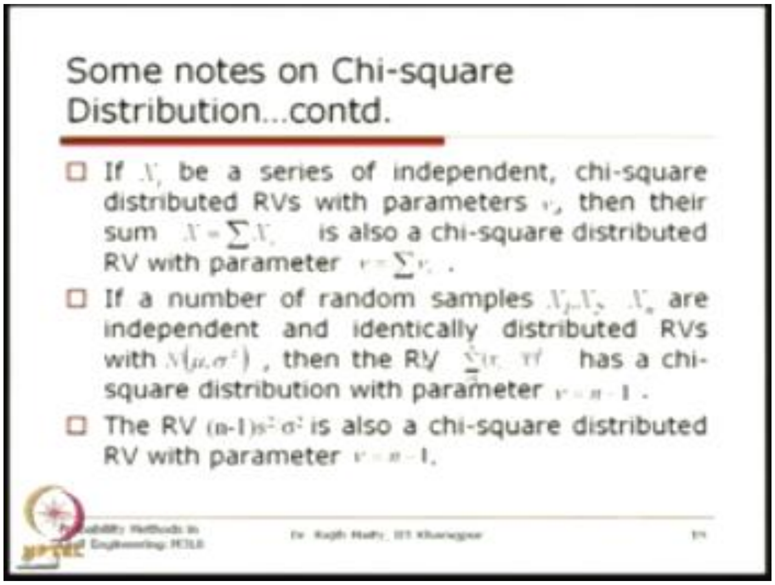
So, while calculating this mean the one degrees of freedom is lost and that is why this, degrees of freedom for this random variable which is chi square distribution this degrees of freedom is n-1. Another random variable if I take that this s, s is now the sample estimate of the standard deviation. 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-,1sothis is also this total quantity is also one random variable, which is also a chi square distribution with parameter nu=n-1with 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 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, the 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 γ.

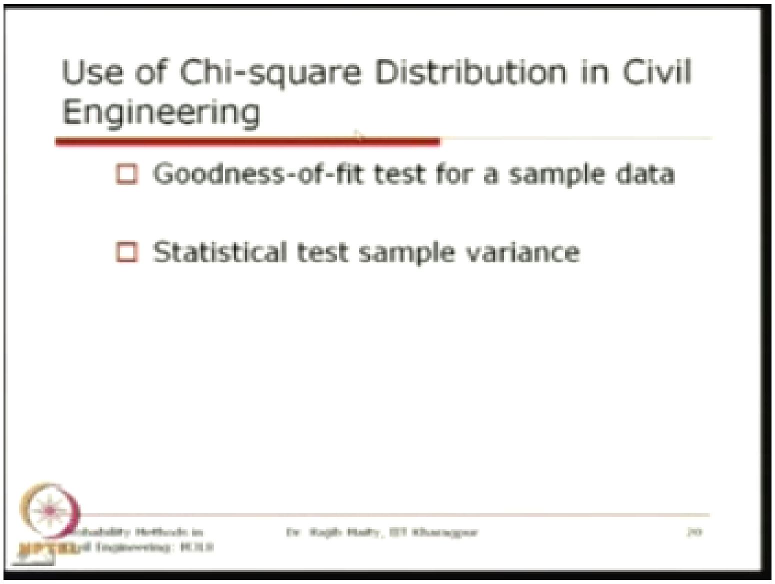
So, that if we take the gamma distribution then what we can do for each point of the sample I can calculate what is the distance of the two cumulative distribution function and that distance, if I just square them of add them up and obviously, then we can see from this discussion that, particular random variable will follow a chi square distribution. So, generally for the goodness of fittest that is why we use the chi square test.

(Refer Slide Time: 39:15)



Secondly the another application is that for this s^2 so when this s^2 that means the sample estimate of the variance so for 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.

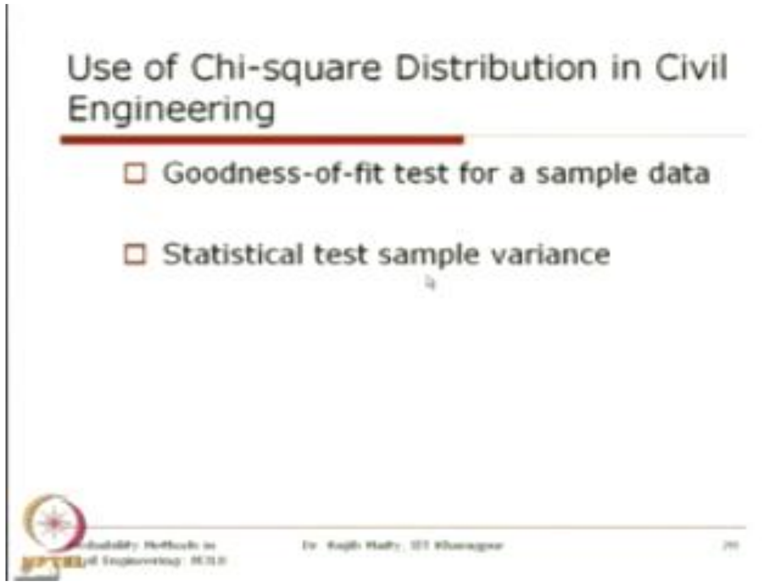
(Refer Slide Time: 39:52)



Well, so first of all I want to do abstract to this one the use of the title of the slide that use of chi square distribution in civil engineering. 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.

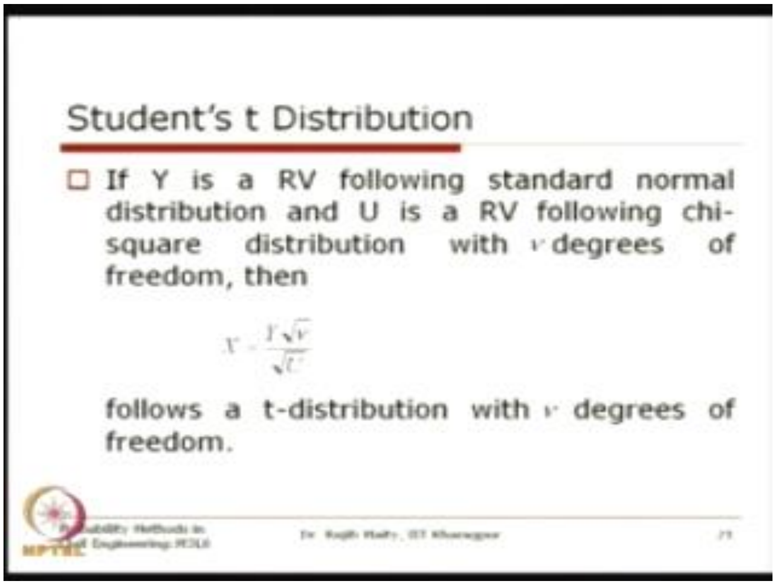
Because these applications are more generally, more general in nature and this is the civil engineering is also not violated. So even though I mention it here that in civil engineering, but this is in general for all the problems in hand. So, this two are we are just now we discuss goodness-of-fit test 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-fittest.

(Refer Slide Time: 41:02)



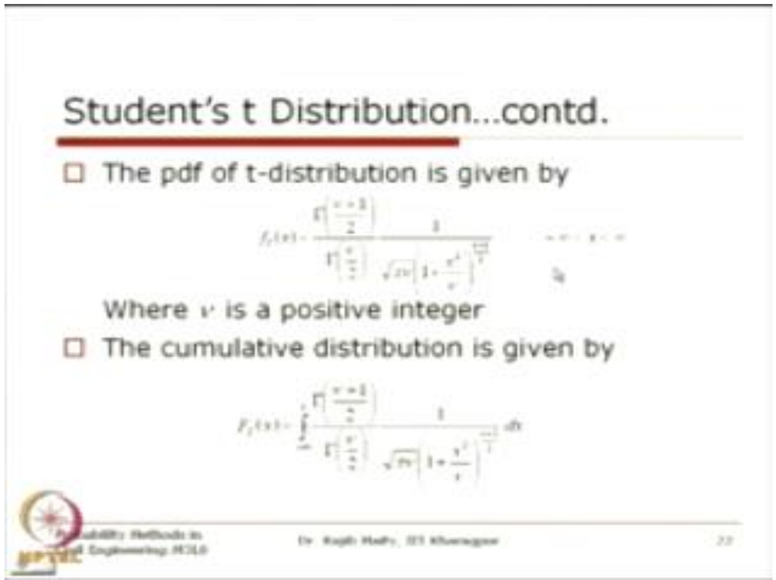
And again the statistical test for the sample variance which also we discuss in this last slide.

(Refer Slide Time: 41:12)



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 nu degrees of freedom? Then if we get another random variable like this that is Y /√ of the chi square distribution normalized by its degrees of freedom, that is U / μ whole square root then, we will get this quantity which is another random variable and this is as then found to follow a student is student t distribution. This is known as t distribution or student t distribution with the degrees of freedom equals to μ.

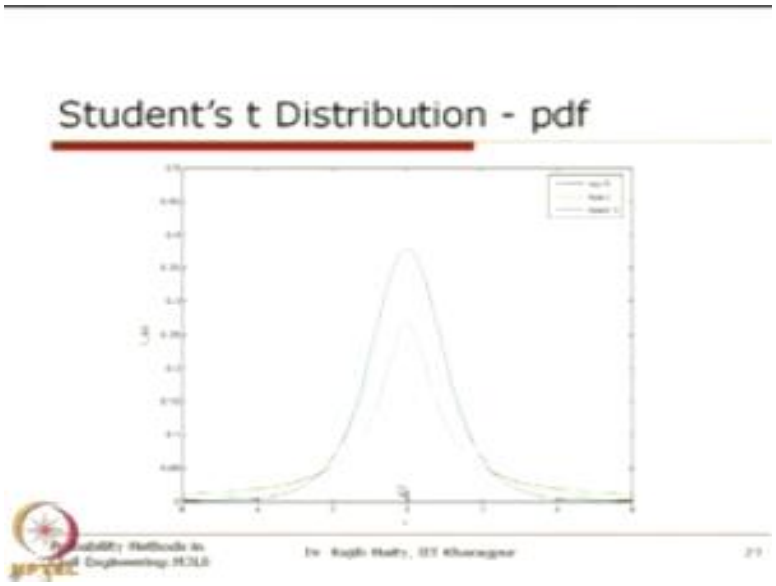
(Refer Slide Time: 42:02)



So its distribution form looks like this the pdf of this t distribution is given by this form that is γ of μ + 1/ 2/ γ μ/2 this is the normalizing constant to make that mixture the total area under this pdf = 1x 1/$\sqrt{2\pi}$ μ x 1 + x^2/ $\mu^{\mu+1}$ /2 and it 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.

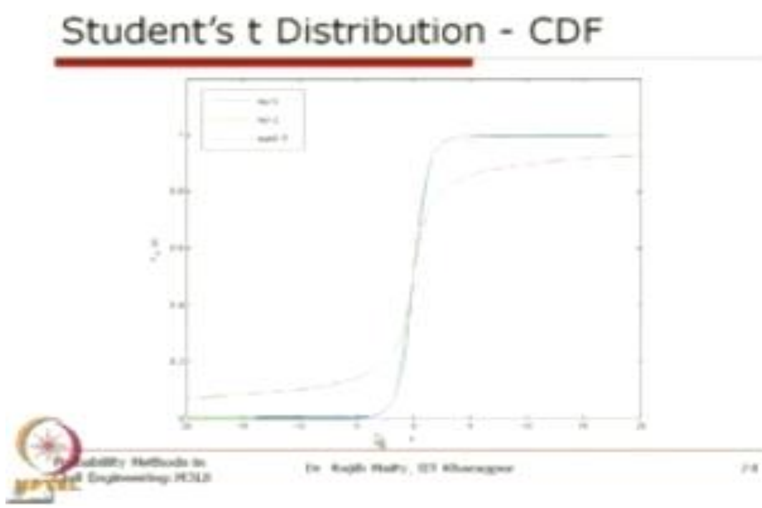
So this μ again is the positive integer which is the degree of freedom for this distribution. And to get this cumulative distribution again it is known that from this left hand support that is minus infinity in this case to specific value x this integration will give you the cumulative distribution of this pdf. These are generally done numerically, because of close form of this integration is not available.

(Refer Slide Time: 43:22)



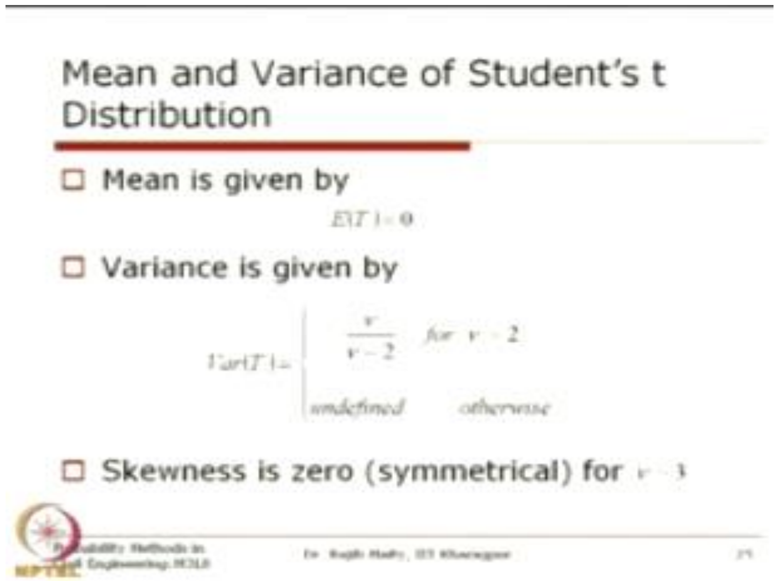
Now again here we have shown different examples of this different values of this nu and you can see that here what we can see is that if we increase the value of this μ this is this increase is peak is increases with the increase of this degrees of freedom and it can be shown that for the high value of this μ, 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: 44:05)



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 these degrees of freedom is concerned.

(Refer Slide Time: 44:37)

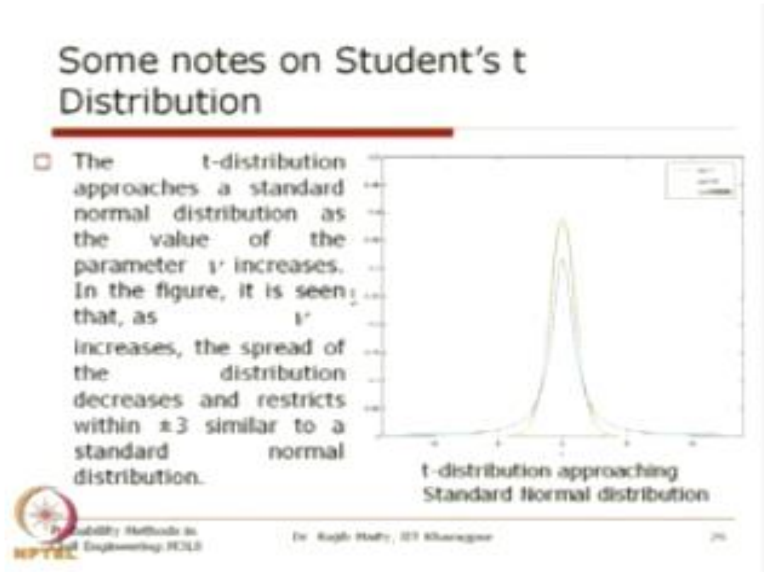


Now these are again the first few moments of this distribution the first of all the mean the mean of the students T distribution is equals to 0 when we are talking about this standard T distribution because this is the standard T distribution when we say. So that is basically is being used when we do 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 μ / μ -2 and this is for when this nu is greater than 2 the degrees of freedom greater than2 otherwise it is undefined. So you can see that this is again so this as this new increases. So this is also become 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 $v < 2$ you know that. So that is why this variance for the student t distribution for $\mu < 2$ is undefined and skewness is zero. We know that we have seen that nature of this pdf. So this skewness is zero as this distribution is symmetric the skewness coefficient if you just see then this is generally defined for this μ > 3, 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)

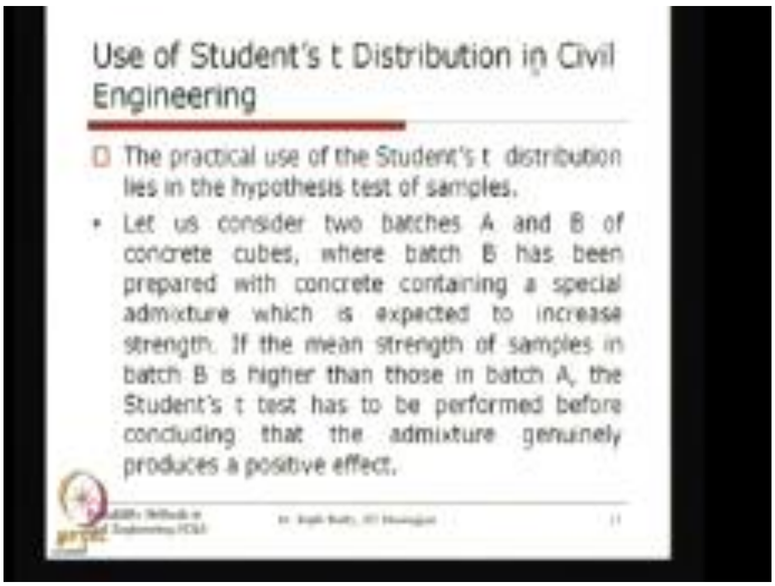


So here if we see that this t distribution, this t distribution approaches to a standard normal distribution as the value of the parameter μ increases. In the figure it is seen that as nu increases the spread of the distribution decreases and restrict within this +-3 to a similar to a standard normal distribution. This is what we are just discussing while giving this example of this the effect of these different values of this μ.

Here also there are three plots are shown: the first is blue one is for the lowest μ here in this case this μ = 1 in this case for this blue line and then and then if we increase that μ that is for this red one you can see that it is μ =10 for this green it is n 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 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 off 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 logicalities case because, for this standard normal distribution effectively the sample size should be infinity.

(Refer Slide Time: 48:45)

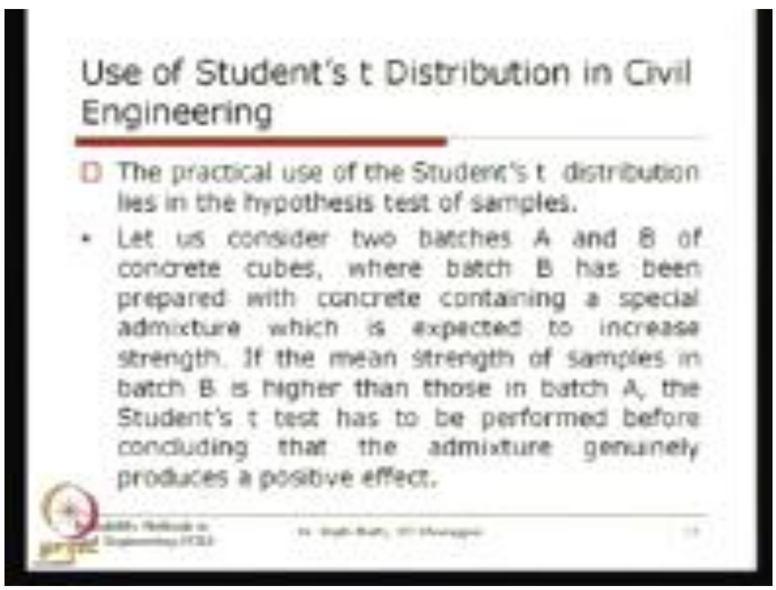


So, use of student t distribution again I am discussing with respect to the civil engineering, but this applications application of all this three distributions student t distribution, chi square and F distribution are more general in nature. The practical use of student t distribution lies in the hypothesis test of this sample particularly, for the samples mean. So, two samples are available to us and if we have the two different mean then the then the target is the weather, we have to estimate that weather the 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 takes 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.

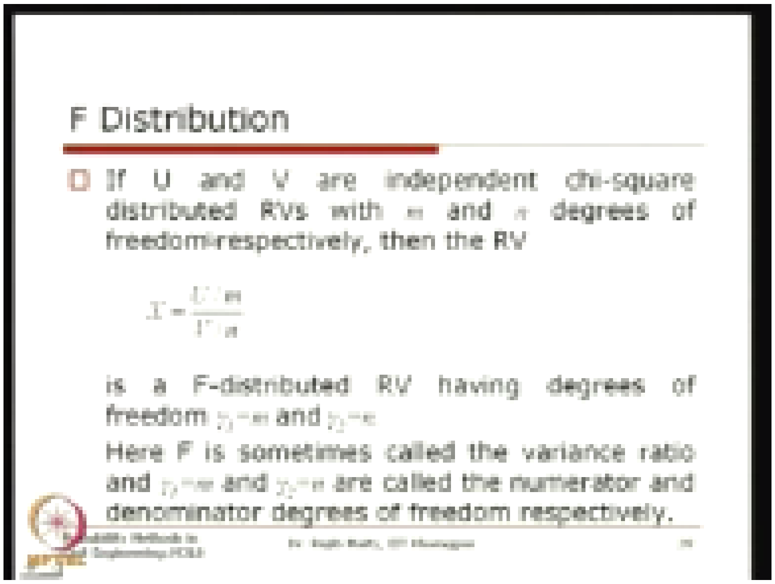
So, in that kind of test what we have to do is that we have to do the two different batches of this concrete cube in one batch that admixture will be added and other batch without that particular ad mixtures. Then we have to take that sample of the calculate that sample mean and then, if you want to 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 hypothesis we have to test while testing this kind of hypothesis where the mean is in involved then we have to follow the student t distribution. This is what is explained here

(Refer Slide Time: 50:37)



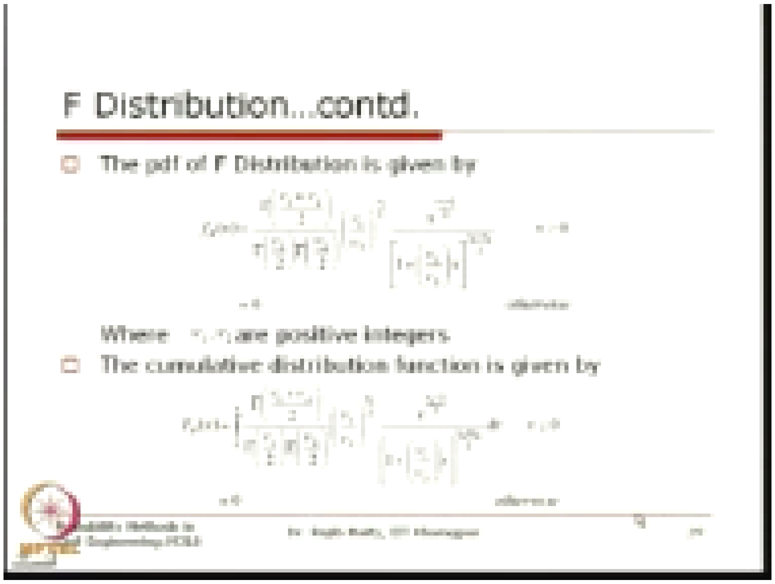
That let us consider two batches A and B of the concrete cube where the batch B has been prepared with the concrete containing a special admixture which is expected to increase the its strength. If the mean strength of the sample in the batch B is higher than those in the batch A the student t test has to be performed e fore 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 in involved we use this student t distribution. The last one in this series is the F distribution.

(Refer Slide Time: 51:26)



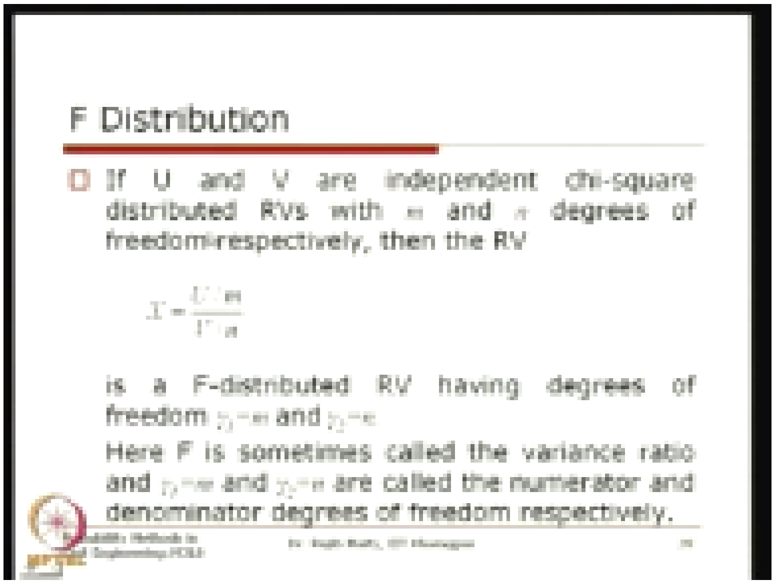
If the u and vary independent chi square distributed with m and n degrees of freedom. Then this their ratio that is U/ m and v / n we will get a new random variable which is F distributed and for the F distribution they are two degrees of freedom and here. So, that is gamma 1 and gamma 2 as denoted is 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 ton are called the numerator and denominator degrees of freedom respectively.

(Refer Slide Time: 52:05)



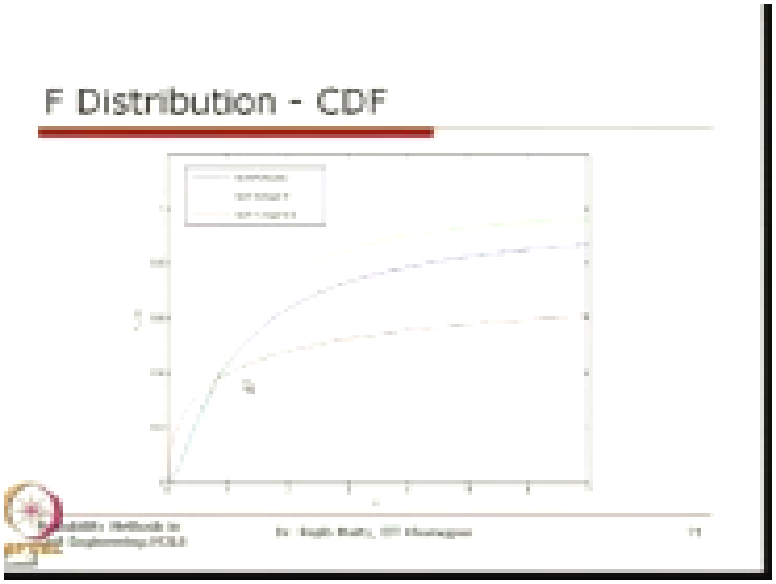
So, this is the general form of this pdf of the F distribution that is gamma nu 1 plus nu 2 by 2 divided by gamma nu 1 by 2 gamma nu 2 by 2 nu 1 by nu 2 power nu 1 by 2 this is the full the normalizing constant, multiplied by x power nu 1 minus 2 by 2 1 plus nu 1 by nu 2 into x power nu 1 plus nu 2 by 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: 52:49)



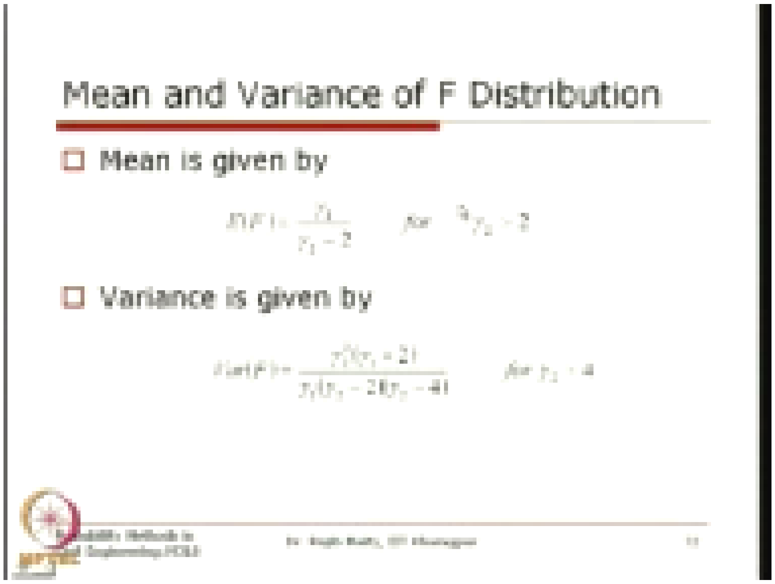
Now, this is the effect of this different values of this combination of two degrees of freedom one is the nu 1and nu 2.We can see that thus red, blue and green different combinations are shown here for different set of combination of degrees of freedom.

(Refer Slide Time: 53:13)



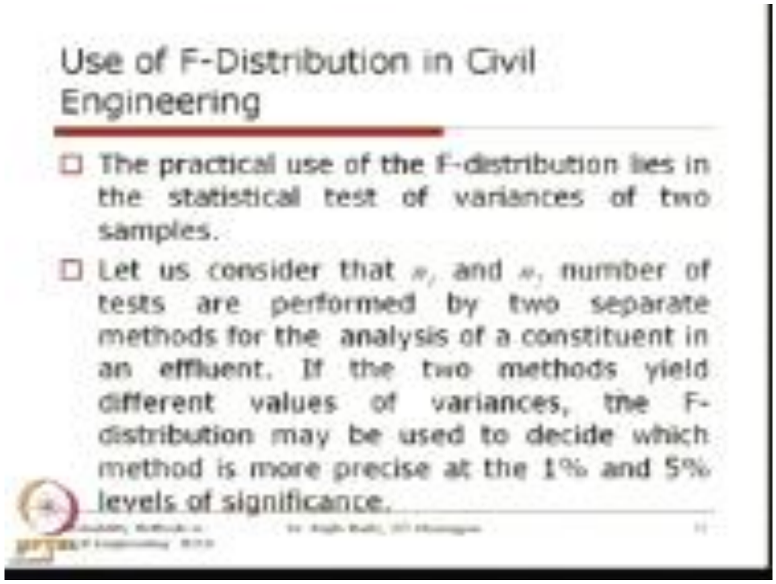
And their respective cumulative distribution functions are shown here which is starting; obviously, from 1 and it is asymptotic to1 starting from 0 going and asymptotic to 1.

(Refer Slide Time: 53:26)



The mean of this F distribution is the this μ_1 by μ_1 μ_2 minus 2 for μ_2 greater than 2 and this variance is defined by this μ_1 square multiplied by a μ_1 plus 2 divided by μ_1 in to μ_2 minus 2 multiplied by μ 2 minus 4 for μ 2 greater than 4.

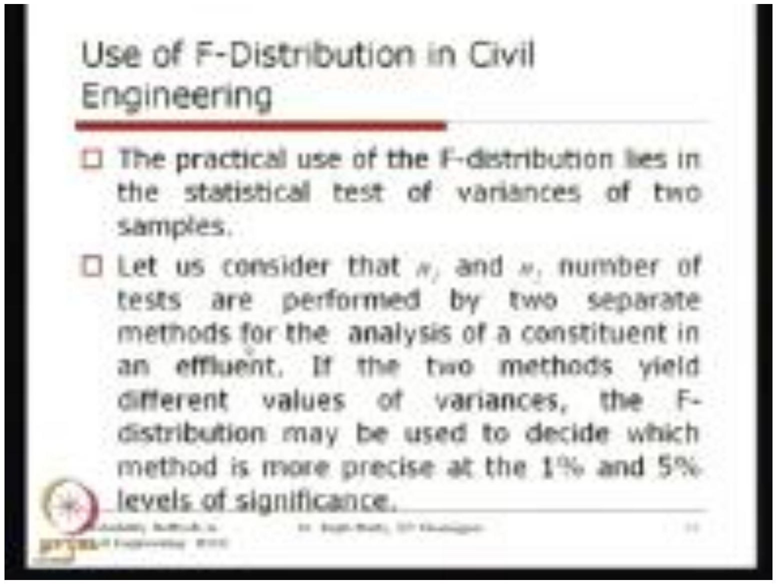
(Refer Slide Time: 53:50)



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 hat 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 ratios should be either greater than 1 or less than1 are equal to 1say if equal to 1 then both the samples, both the samples' standard deviations 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 populations follow a F distribution.

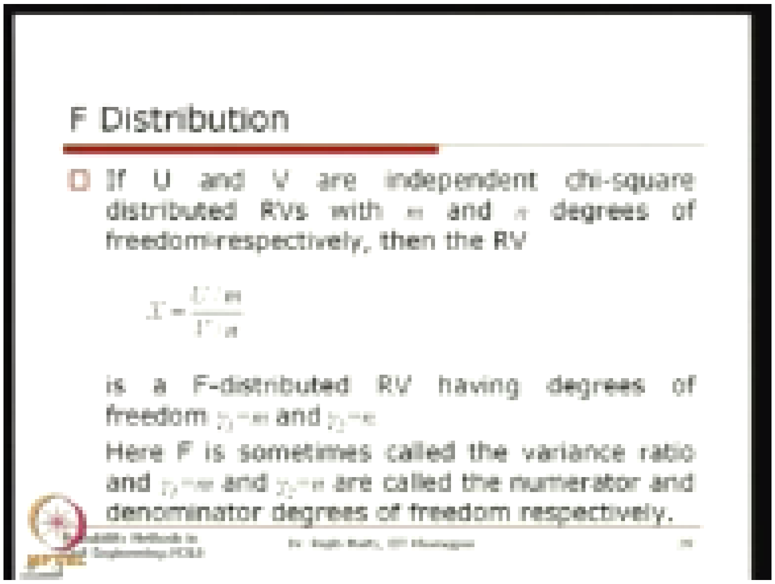
(Refer Slide Time: 55:03)



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 n1 and n2 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 1percent 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 substituent module and that last modules even when we discuss about the statistical test.

(Refer Slide Time: 55:46)



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 is the mostly used into address the different problems in civil engineering. Thank you.

Probability Methods in Civil Engineering

End of Lecture 13

Next: "Functions of Single Random Variables"
In Lec 14

NPTEL Video Recording Team
NPTEL Web Editing Team

Technical Superintendents

Computer Technicians

A IIT Kharagpur Production

www.nptel.iitm.ac.in