Probability Methods in civil Engineering

Prof. RajibMaithy
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

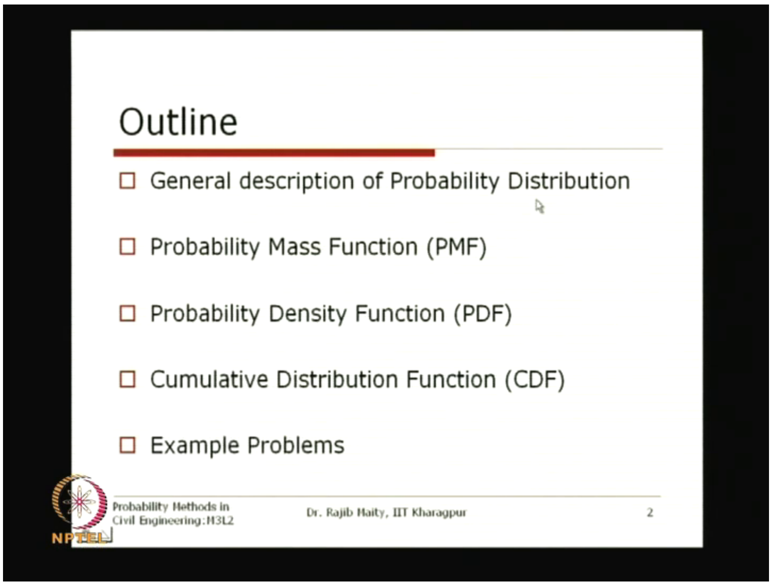
Module # 03

Lecture # 07

Probability distribution of random variables

Hello there.Welcome to the second lecture of module three. In this lecture, we will know about the probability distribution of random variable.In the last lecture, we have seen that definition and concept of this random variable.Now, this concept and definition of this random variable, generally is useful in this probability theorythrough its probability distribution.We have to know that on top of the specific range, over the specified range of one particular random variable, how its probabilities are distributed.This is what we will discuss in today's class.

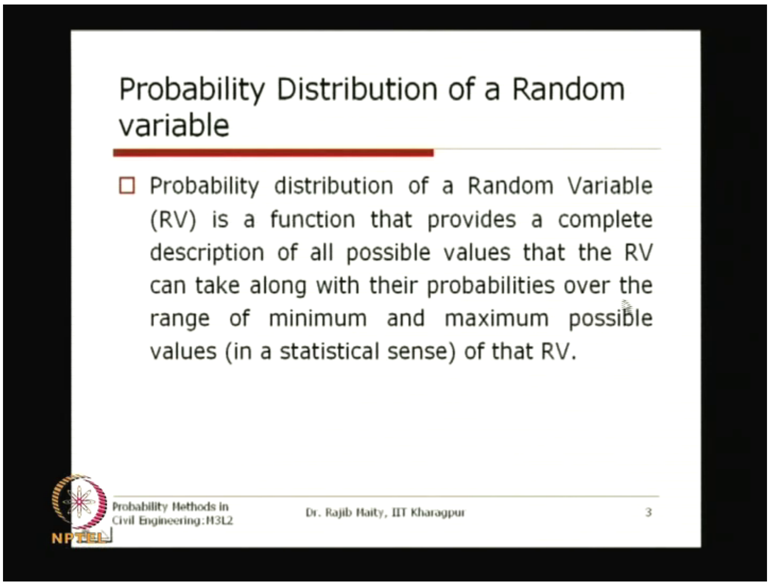
(Refer Slide Time: 01:16)



So our outline for today's presentation is, first we will discuss about the general description of the probability distribution; how this probability distribution, what is, what it is all about, how we can define the probability distribution for a random variable.

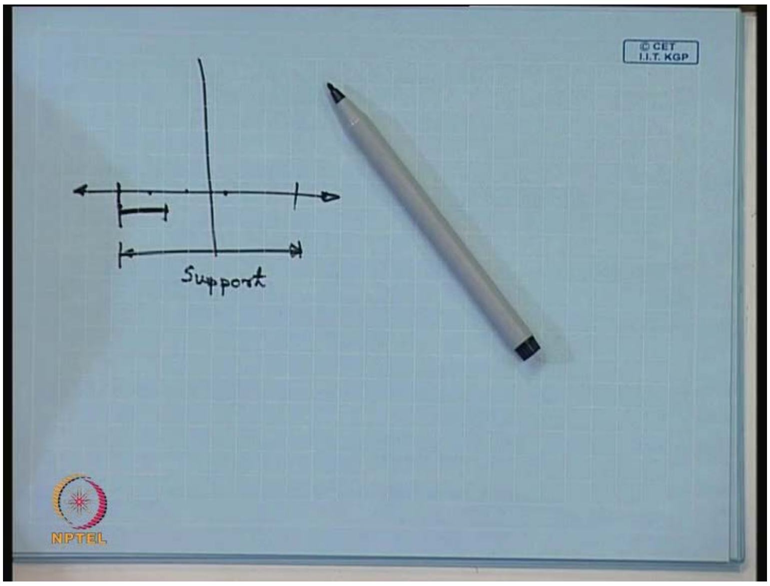
Basically, there are two different types of random variables we will consider.One is thatprobability;one is that discrete random variable and then the continuous random variable.This probability mass function will be discussed for this, which is for discrete random variable and probability density function is for this continuous random variable and their cumulative distributions is also known as that cumulative distribution function.So this will be discussed and for all these things, we will see some example problems as well.

(Refer Slide Time: 02:05)



Probability distribution of a random variable, it says probability distribution of a random variable is a function that provides a complete description of all possible values that the random variable can take along with their probabilities over the range of minimum and maximum possible values in a statistical sense of that random variable.So here, the meaning is that a random variable, in the last class we have seen that this random variable can take some specific value over the, for one particular random experiment the specified that sample space can be correspondence to the real length through the random variable which is nothing but the random variable and that random variable, is generally a functional correspondence to the real length some numbers.So, it can take some take some numbers that is a generally having certain range.

(Refer Slide time: 03:15)



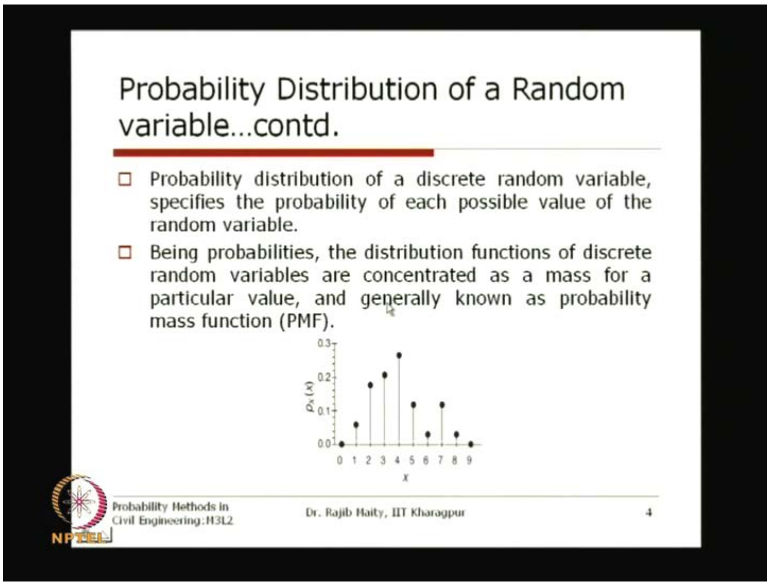
And over this range, if we have just seen it here, over this range of this variables, how, suppose that one, suppose that one random variable, this is yours, this is your the axis for this random variable. If you just see this side is your probability, then if I say that this is the range of this random variable. Now thing is that, here how this for each region if it is if it is continuous, then for some region, how this probability is distributed over it, for this entire entire range of this random variable.

In the context of this probability probability distribution, this is known as the support of support of the random variable.Support of the random variable andover this, if it is a discrete random variable, then for this specific values the distribution, specific values the probability will be specified.On the other hand, if it is continuous, then it will be it will be distributed as a function over this entire support.

Thus, now again,the another point here is the maximum and minimum possible values.This maximum and minimum possible values, in obviously, this is in term, in the statistical sense. What is meant is that, may be a random sampling if you take, for any random variable, if you just see for one observation, that the maximum of that oneor minimum of the sample need not be the maximum or minimum of that random variable.So, that can have some other values that will, that is what it is meant by this statistical sense.This will be obviously discussed again in details when we are going to some specific distribution.

For the time being, what is important is that, a random variable it is having a specified range and this probability distribution gives us gives us the distribution of the probability specified for each, for all possible values that the random variable can take.

(Refer Slide Time: 05:30)



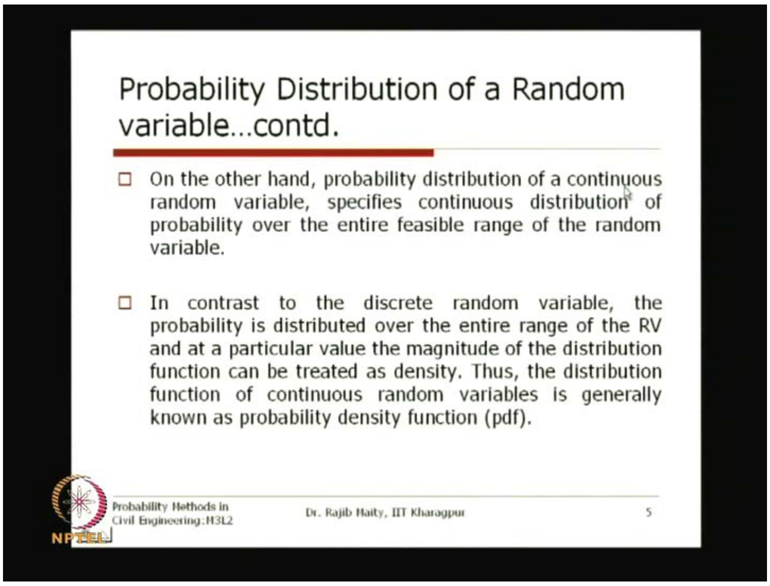
Now, as as we just mentioning, that it can, so the two different concept;one is for this discrete random variable another one is for the continuous random variable.As we discussed in the last class, that discrete random variable here means that it can take some specific value over the range of this random random variable.It cannot take all possible values, even though, traditionally or in most of the cases, this this discrete random variable takes the integer value.But, that is only the concept.It can take any specificvalue not only integers but, it takes only that specific value.So, that is a discrete random variable.On the other hand, the continuous random variable, can take any value over the inter support of that, of the distribution or inter range of that random variable.

So in, so first we will discuss the probability distribution. We will discuss with respect to this discrete random variable.It states that probability distribution of a discrete random variable specifies the probability of each possible value of the random random variable.So can see here, I, there is one random variable x, which can take the values 0 1 2 and so on up to 9.So at each and every possible values that the random variable can take, the probabilities defined there.

So now, so there is nothing in between two integers because,this random variable take the integer integer values only.So, in between two integer values say for example, between this one and in between this two, there is nothing is specified here.So, this space is entirely entirely is not specified by this by this distribution.So, here what we can say that, this particular, at a particular point of this random variable, for a particular specific value, this one can be treated as a mass.So, this is concentrated at a particular point.So, this is why, so this is can be treated as a concentrated mass.That is why it states, that being probabilities, the distribution function of the discrete random variable are concentrated as a as a mass for the for a particular value and that is why it is generally known as the probability mass function and abbreviated as probability pmf.

So, what is meant here, that as there is nothing specified in between two specific values of the random variable. So, what is specified for this specific value of the random variable is can be treated as a mass of probability That can be treated as a mass, so that is why, this kind of distribution, we know it is known as the probability mass function, abbreviated as pmf.

(Refer Slide Time: 08:40)



On the other hand, as for the for the continuous random variable, this is not the case.So, this can be, this is specified over the, over a over a region.So, on the other hand, the probability distribution of a continuous random variable specifies, continuous distribution of the probability over the entire feasible range of this of the random variable.

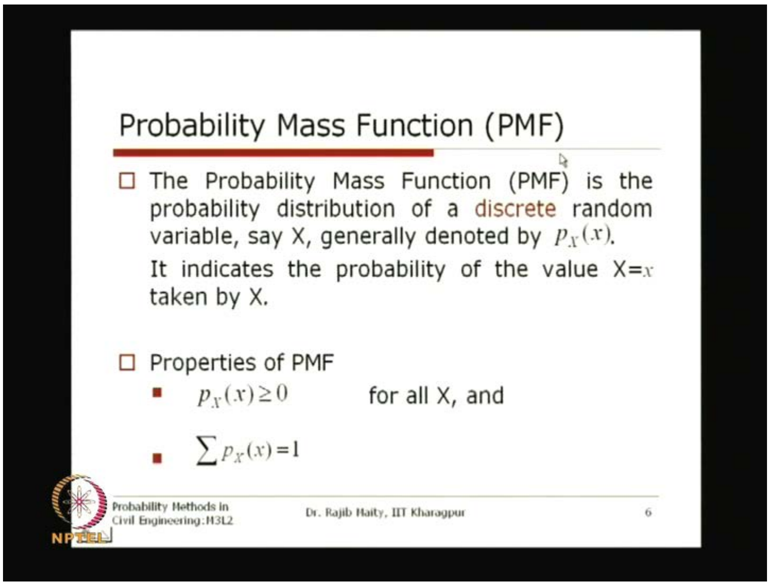
So, if the random variable is continuous, then that random variable is specified over a region, over a range of over a range.So, the distribution function should be specified in terms of a, obviously, in terms of a function over the entire range.And, in contrast to the discrete random variable, the probability is distributed over the entire range of that random variable and at a particular value the magnitude of the distribution function can be treated as density.Thus, the distribution function of continuous random variablesis generally known as probability density function and abbreviated as lower case of pdf.

So, this point, that is the, its concept of the density,I will just explain it here.As, I was telling that, if this is the entire support of the,so now what I am discussing this is in case of the continuous random random variable. So, for this continuous random variable, if I say that, fine, this distribution is specified distribution is shown like this.So, for what it, first of all, what is what is showing that for some region, this this probability is lower than with compared to some other region.For example, this region here, it is more and this region here, it is less.Now, thing is that, if in this case, if I just mentioned that a specific value of that random variable x, so this, is that specific value.Now, what does this this implies is, this implies the probability.Now, if I just draw the, draw the same thing for the discrete, then, what I have seen just now, what we have seen just now, for a specific value of this random variable, the probability, what is specified; what is concentrated, yes, this nothing but, for the probability.

In between there is nothing is specified.So, in between this region, nothing is specified.But, whatever is specified that is nothing but, the probability.Now, for this one, if I just say for a specific value, what is this height meant.This is very important to know that, this height is not the probability.Then, where is the probability here, so here.The probability means that, so I have to specify a small small region around this, some small region.Then, what we are getting, we are getting some area and that area is nothing but, real probability is showing here.So, at a particular point, if I consider here, this height is nothing but, we can treat that this has a density.The density of the probability here.Once, you are multiplying the density for a for a normal physical science, if you multiply that density with with its mass, then you will get it over.So, similarly, here this is nothing but, the density.If you multiply over a certain range, then we will get the area and that area nothing but, your probability.So, that is why for this distribution, for the continuous random variable, this distribution is the probability density function.

So, this is why the,what density comes here.So, once again, if we just read it in contrast to the discrete random variable, the probability is distributed over the entireentire range of this random variable and at a particular value, the magnitude of the distribution function can be treated as density.Thus, the distribution function is known as this probability density function.This we will be more clear when we are talking about that one of the axiom that we have seen in the earlier earlier classes, that the total probability, obviously, should be close to 1.So, this will be this will be clear in a minute.So here, what we are trying to say, why this word density.

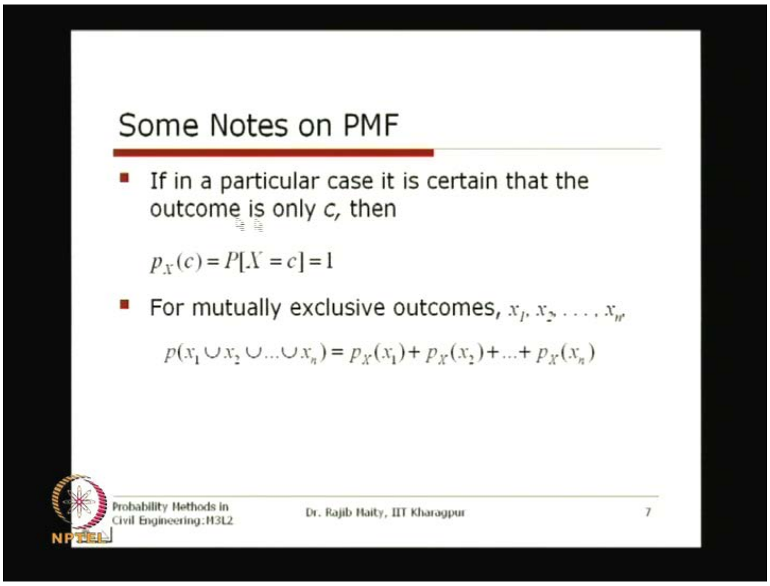
(Refer Slide Time: 13:24)



So next, we will first start with this probability mass, p m f, which is for this, we just now have seen, this is for the discrete random variable.So, whenever we mention that pmf, this is for the discrete random variable.The probability mass function,pmf, is the probability distribution of a discrete random variable, say x, generally this is denoted by p x x.So, here this notation is important.This p is the lower case; this p is lower case; this subscript x is the, is denoting the random variable.So, this function is for which random variable, this is shown as the upper case letter as a subscript to this one. And this one is lower case, which is nothing but a specific value of this one.This we also discussed in the context of some other distribution in the last lecture.

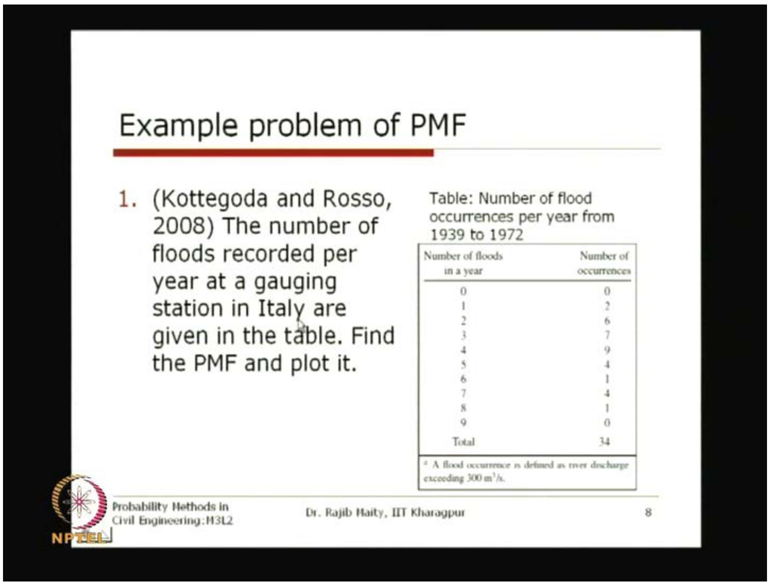
That so this one, this lower case is the specific value of the random variable, which is shown here, x and this small p, this is nothing but, for this probability mass function.Now, it indicates the probability of the value x equals to that specific value x taken by the random variable x.So, there are some properties for this random variable, just to what we are just telling, indicating in this last slide is that, the first property is that for each and every value, whatever the this random variable can take should be greater than equal to 0.So, this probability can never be negative.So, this is an non negative number and the summation of all this all this probabilities, now this x is defined over some specific values of this one, that is, so they are from specific values of x where this probability is defined.Now, if we add up for all this all this possible value, all the possible values of this x, the probability of the all possible values of x, if you add up, then it should end up to 1.So these two are the properties of this probability mass function.

(Refer Slide Time: 15:40)



Now, some notes on this pmf.This is, these are obvious,but still it is important to important to mention here, that in a particular case it is, if it is certain that the outcome is only c.So, for a random variable and saying that there is a only one outcome and that outcome is certain and that outcome is c.Then, what is this p x c?The distribution function is nothing but, p x c which implies the probability of x is equals to c.This is the only outcome that is that is feasible and this is certain outcome, so this will obviously come up.So, this is equals, so this should entirely be equal to 1 to satisfy all this properties of this pmf.

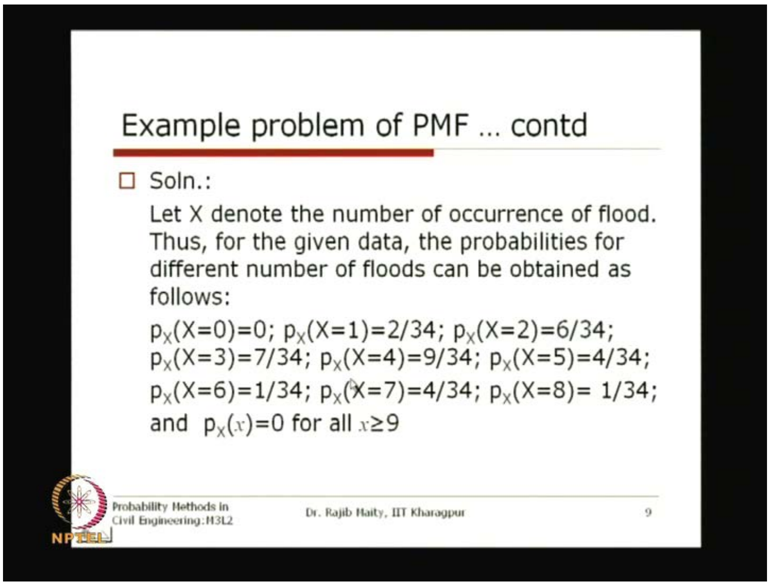
On the other hand, if there are some mutually exclusive outcomes for one random variable, that is x 1 x 2 up to xn.Now, if we just say thesethis values, this specific values of a random variable x are mutually exclusive. Then, if you want to calculate, what is the probability of either of this mutually exclusive exclusive value that the random variable contains should be equal to the summation of their individual individual probabilities.So, this is obvious in in case of a throwing a dice.There are six possible outcomes and if you say that all the outcomes are equally feasible, equally equallyequally possible and if I just take that number 1 2 to number 4, then the probability, the total probability that the that the random variable random variable will take either 1 or 2 or 3 or 4 will be equal to the summation of the probability of getting 1 plus summation of the probability plus probability of getting 2 plus probability of getting 3 and plus probability of getting 4.So, we know that this 1 to 4, these events and mutually exclusive, so this can be, so this properties we explained earlier in the context of the random variable.Here, we are explaining in the context of the specific value, that a discrete random variable can take. (Refer Slide Time: 17:55)



So, one small example will take on this on this probability mass mass function.This is taken from the Kottegoda and Rosso book.Kottegoda and Rosso bookThe number of floods recorded per year at a gauging station in Italy are given in this table.Find the probability mass function and plot it.Now, here for this 1939 to 1972, there are so many number of floods are noted here.So, 0 floods, this table is this 0 flood has occurred in 0 years.So, 1 flood has occurred in 2 years, 2 floodshas occurred in 6 years, 3 floods has occurred in 7 years, in this way.So, the total number, 9 floods in a year occur for 0 occurrence.So, if you just add of the total years, this 34 should match with the, whatever the data that is available to us.Now, we have to define the pmf for this one. So, this kind of problem, the first thing that that we should think that, what is the random variable that we are talking about?

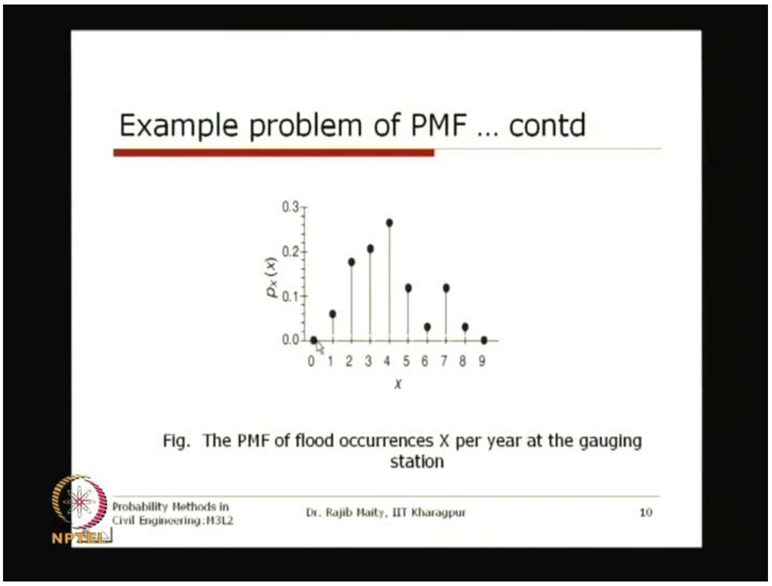
So, this this is one of the, even though this is sometime for this kind of problem, this is quite obvious but, this is important to know that what is the random variable random variable here.So,occurring the flood is not the random variable. I repeat, the occurring of a flood event is not the random variable here, rather the number of floods in a year that is the random random variable.

(Refer Slide Time: 19:23)



So, while solving this problem, first of all, what we will do, we will first define that, what is the random variable that we are mentioning.So, let x, that is the random variable denote the number of occurrence of flood.Thus, for the given data, the probabilities of the different number of the floods can be obtained as follows.So, px equals to 0 number, if we take this number 0, that is that is the number of flood occurring 0 flood, no flood in a year should be equals to 0.Because, from this table, we see there is no such a year where the number of flood is 0.Similarly, if we want to know the, what is the probability that x equals to 1.That means, from this sample data, obviously, so x equals to 1.So, that means there are two such occurrences out of 34 years.So, 2 by 34 should be the probability for that specific value of the random variable, that is x equals to 1.

(Refer Slide Time: 21:31)

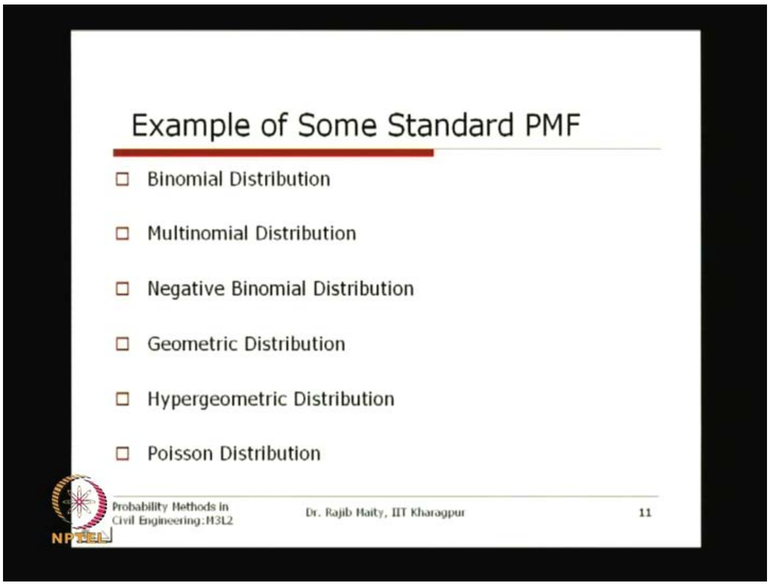


So, this is what the probability of x equals to 1 is equals to 2 by 34.Similarly, for x equals to 2, 6 by 34, x equals to 3, 7 by 34 and in soon we are going and getting all this probability values.Now, similarly, we are assuming one thing, that as, so number of flood, 9 floods in a in a year, this is occurring 0 and all we just got the summation this one.So, there is no other occur numbers of flood.If, I just take 10 11, there are also number of occurrence in a year, 10 floods in a year 0,11 floods in a year 0.

So, to complete the the definition of this pmf, x x equals to 0 for all x greater than 9.Now, simply we have to plot this one as a mass mass function for this specific values only.Nothing in between 1 and 2 because, that is not specified, which is obvious for the discrete random variable.So, this is the plot.So for the 0, the probability is 0 and obviously for 9, probability is 0, for all higher values is also 0.For 1, these are the probabilities, whatever the value we got here, so we are getting this masses.Now, obviously do not confuse that this, what is the meaning of this line.Basically, this line, this solid line has no meaning; just for as a geometric reference that this point refers to this number 1.Otherwise, a simple a single dot at this point should be sufficient to display the probability mass function.

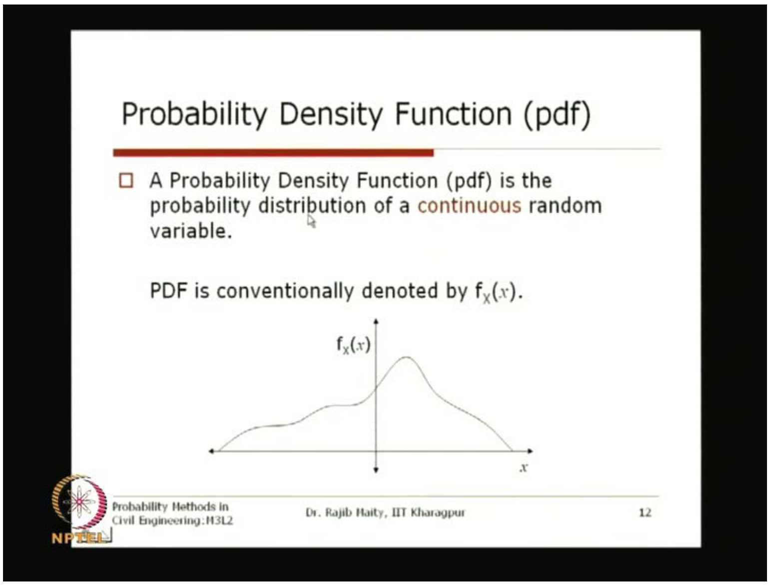
Now, we can, from this probability mass function, we got the probability mass function for the data that we have.Several things can be answered from this one.If it is asked that, what is the probability that number of flood is greater than equal to 5?So, if I say that number of number of floods greater than equal 5, then obviously, I have to just add up these values.If I add up this values, for the probability for 5 6 7 8 and 9, then I will get what is the probability that the number of flood in a year is greater than 5.So, this is the utility of this pmf, that all of this kind of answer we will get from this probability mass function.This we will again see while we are discussing the cumulative distribution function.

(Refer Slide Time: 22:57)



So, this is for the discrete and there are some standard example.There are some standard probability mass function, that is for the discrete random variable, this is the binomial distribution.We will discuss all this distribution again in detail in the in the successive lectures.For the time being, we can just know the names.Binomial distribution, that means there are binomial distribution means, there are, this is a bernoulli trial where there are two possible outcomes.One, we just we just tell it a successes and the success for each trail is predefined, which is known.So, now the number of, number success out of n successive such trails, that is the number, so that number is a random variable and that random variable follow this binomial distribution.Similarly, if there are more than two, more than two outcomes, then for how many success we are getting in a set of, say 1 tok and all this success rates are known, then the vector, that is that x 1 x 2 x 3 up to x k, we will follow the multinomial distribution.Similarly, there are different definition for this negative binomial distribution, geometric distribution,hypergeometric distribution, poisson distribution.These are example of this, the distribution of the, distribution of discrete random variable which will be covered in the successive lectures.Now,we will go to the distribution function of that continuous random variable.

(Refer Slide Time: 24:35)

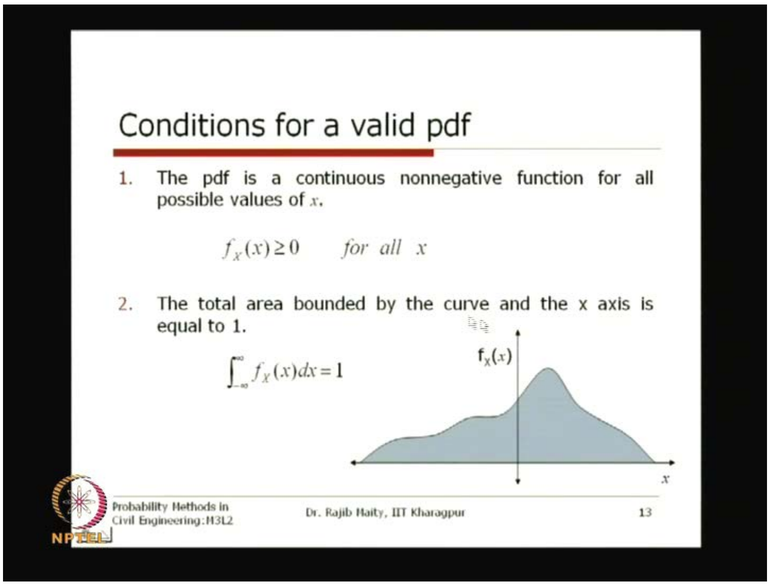


So, this continuous random variable,when we are taking, we call it as the probability density function.Why it is density?Just we discussed now, so a probability density function abbreviated as lower case of pdf is the probability distribution of a continuous random random variable.Do not confuse about this abbreviated form.This is, these are the abbreviation will be followed for this lecture.But, in some standard reference book, you may get some other notation.But here, we have to mean that probability density function, we generally we generally abbreviated as lower case of pdf just to differentiated from this cumulative distribution function, where the d stands for the distribution.Here this d stands for the density.

So, this pdf is the probability distribution of continuous random variable.Generally, it is denoted by this f x x, which we discussed last time also, that this is the random variable which is the upper case letter and this is the specific value of the random variable which is shown as the lower case of this of the letter.So, here again, you can see that this is the distribution function defined over this one. So, this is the total range of this random variable of that.It can take and obviously here, the density is more and here the density is less.

Now, we will see that there are obviously, it is not that any function I will take and I can tell that this is the probability density function.That is not the case.There are certain conditions should be should be followed to make a particular function to be a probability density function and those conditions are this.

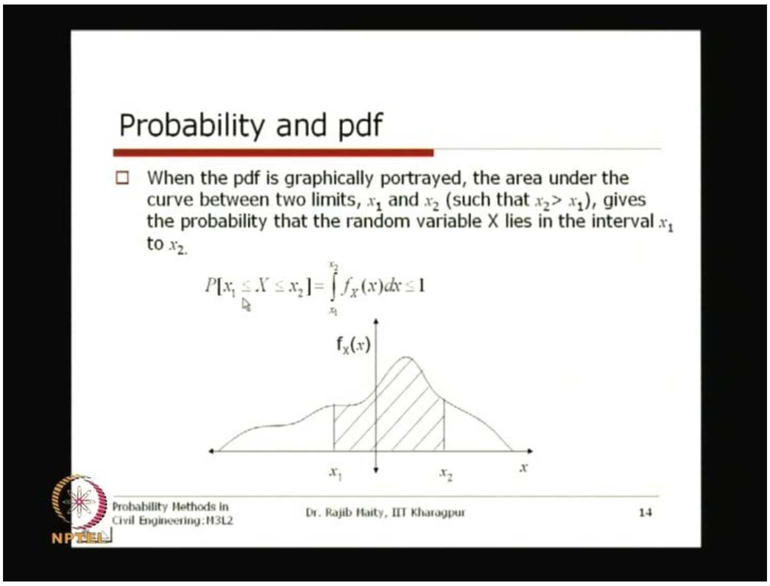
(Refer Slide Time: 26:26)



So, there are two condition conditions for a valid pdf.The pdfthe p d f is a continuous nonnegative function for all possible values of x.So f x x, for all x should be greater than equal to 0.This is basically coming from the first axiom of the probability.So, in this graph, so everything that is coming above this, towards the positive y axis and the total area bounded by the curve and the x axis is equals to 1.

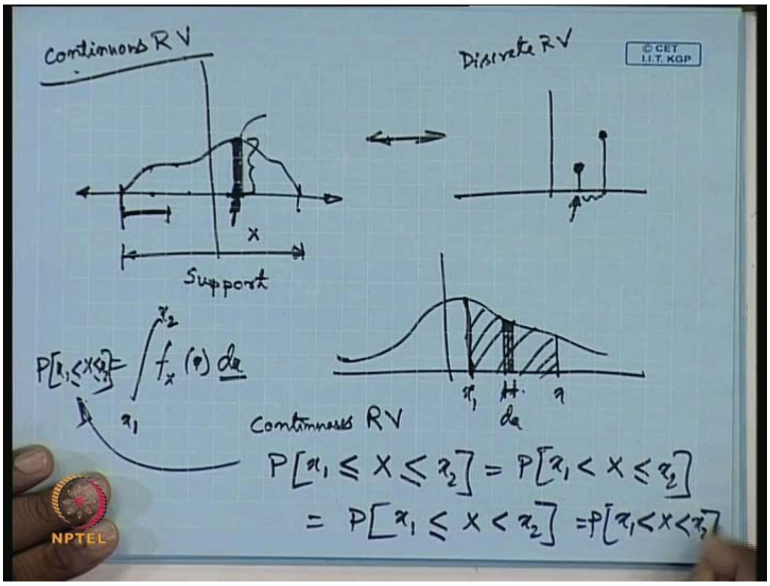
So that, this shaded area, what you see here, this shaded area below this graph, above this, above this axis above this axis should be equals to 1.So, this is basically mean that, if I take the inter range of this random variable, then this is becoming a certain event.So, any one, any possible value will take here.So, this is the entire set of the sample space and which is equals to 1.So, the total probability of this entire end should be equals to 1.So here, we have just written the minus infinity plus infinity of the x should take care about this full one. So, obviously this is reducing to this, the lower limit and the upper limit of this one, because the restof the places, this random variable is defined to be 0.So, this is the second condition. So, if any function that is that passes through this two, then that condition can be a valid pdf.

(Refer Slide Time: 28:15)



Now, this is one important concept, what I, what was just discussing, while discussing the density, that when the pdf is graphically portrayed, the area under the curve between two limits, x 1 and x 2 such that, say x 2 is greater than x 1, gives the probability that the random variable x lies in the interval of x 1 to x 2.So, this probability that this random variable will be in between x 2 and x 2 x 1.So, this is graphically nothing but, as we have just telling that, each and every point here, this is this is implied that, this is implied the density for that particular value.So, if I want to know what is the probability that this random variable be within this limit.This is nothing but this hatched area over this in this graph. So, this area will, what how will you get it at this, so we will integrate from this x 1 to x 2.This integration and this will obviously be less than equals to 1, because we know that this total area is equals to 1.Now, again if you just see this integral form, then it looks like this.

(Refer Slide Time: 29:40)

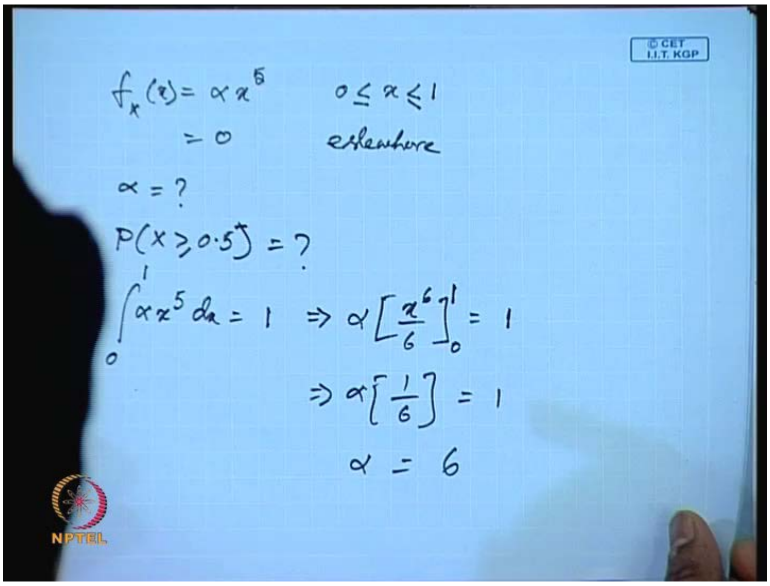


So, the integration here is, so integration what we have seen here is that f x x d xfrom x 1 to x 2.Now, what we have just discussing here, that this is this is the density here.Basically, what we are doing here for this graph, so we are taking a small, so d x is your small strip here.

D x is the small strip here. So, this is your d x and the thisone, this f x gives you that particular value for this area.So, if you multiply these two, which is nothing but the probability, you are getting over the area d x. And, this area, you are just basically adding up from this x 1 to up to x 2.So, that is why you are getting the total area below this below these two limits from x 1 to x 2 and that is nothing but, which is gives you the gives the probability.This this this probability is obviously from this x 1 to x 2.Now one thing, here again is important, so far as the continuous random variable, so highest this way. So far as the continuous random variable is concerned, basically, this probability that x 1 less than equal to, so this sign I am just stressing the point that equal to sign, having this equal to sign or not having this equal to sign, does not mean anything because, ultimately for a particular for a particular specific value the probability is 0 as this range for this particular value over which is the probability is defining is 0.So, this equality sign, inclusion of this equality signor not inclusion of this one does not change the total probability.So, what we can express is that less than equals to x 2 is equals to probability of x 1 less than x less than equals to x 2 and whatever the combination possible is that x 1 less than equals to x less than x 2 equals to probability x 1 less than x less than x 2.

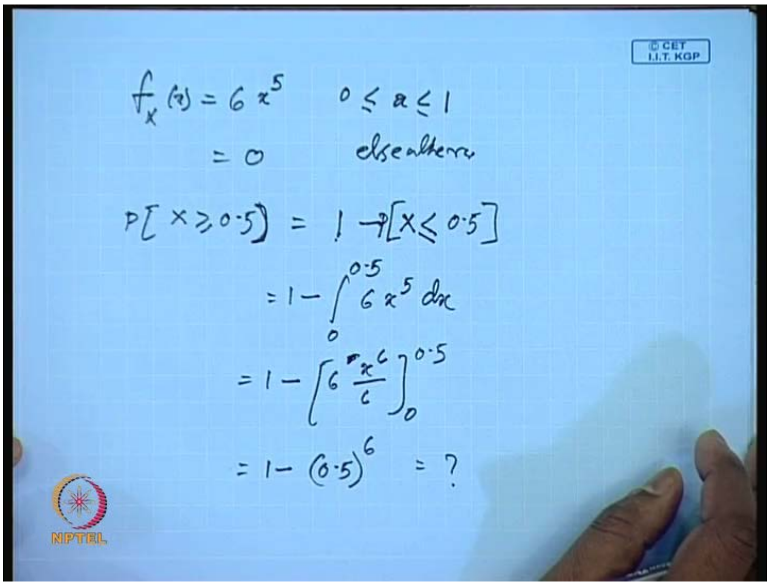
So all these four cases, the probability is same as long as this random variable continuous.This is one important concept here, while you are calculating the probability from the pdf for a continuous random variable.Now, we will take one small example, mathematical example, rather to discuss about this pdf, how to satisfy their properties.

(Refer Slide Time: 32:40)



Suppose that, the function f x x is equals to its alpha x power 5 and which is defined over the zone, for this x value from 0 to 1 and it is 0 elsewhere.Now, to be, this is a valid pdf, what is the value of alpha?That we have to determine and what is the probability of x that is greater than equal to 0.5.

(Refer Slide Time: 34:31)

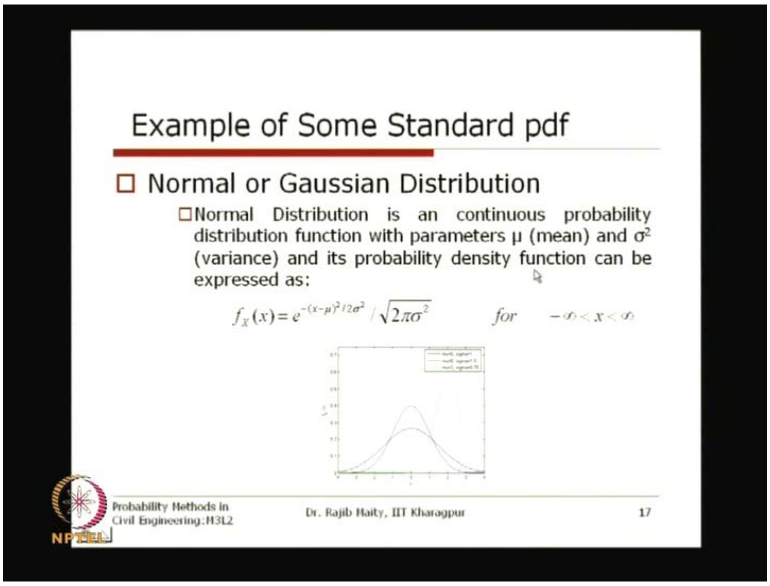


So, if you want to know this first one, that is the, what is the value of this alpha, then we will know that the property that this should be from this entire range of this of this random variable, that is 0 to 1 in this case, that this function should be equals to 1.Now, if we do this one, then this will be alpha, this x power 6 by 6, which is equals to your 1.This is alpha; so, it is 0 to 1 by 6 minus 0 equals to 1, where the alpha is equals to 6.

This value we got.So now, what we will see, that this, so what we got that f x of this x equals to your 6 x power 5 over the range, this x 1 and equals to 0 elsewhere.Now, the second thing is that, what is the probability that x is greater than equal to 0.5.This we can express, that you know that 1 minus probability of x less than or less than equals to, you know that for this continuous random variable, these two are same, is equals to 0.5 less than equal to, so 1 minus this.We can do that from 0 to 0.5 6 x power 5 d x, which is 1 minus 6 by 6.We can just write to here, 0 to 0.5, so it is equals to 1 minus, say 0.5 power 6.So, then we can calculate this one, this probability with the help of this.

For two things we want to discuss here.Now, one is that just to gauge this x is greater than equals to 0.5.What we can do is that, we can just do the integration from 0.5 to 1, because this is the range 0.5 to 1.We can do this integration directly to this function and we can get the probability and answer of it will be the same.What instead of that also what we can do, we know that the total probability is equals to 1.So, 1 minus the rest of this part, from this means here 0 to 0.5.What is the area that we have deducted to get this probability?Basically, this relates its link, to the to the C D Fbecause, we will just see in a minute that what is a C D F. So, from the C D F, we can directly calculate, what is its probability and that probability value, we can put in this place.Instead of, so, that is why it is it is it is replaced in terms of this one, just for one illustration purpose which can be linked to the C D F that we are going to discuss in a minute.But, so far as this particular problem is concerned, we can also calculate the integration from 0.5 to 1 to get this particular probability answer.

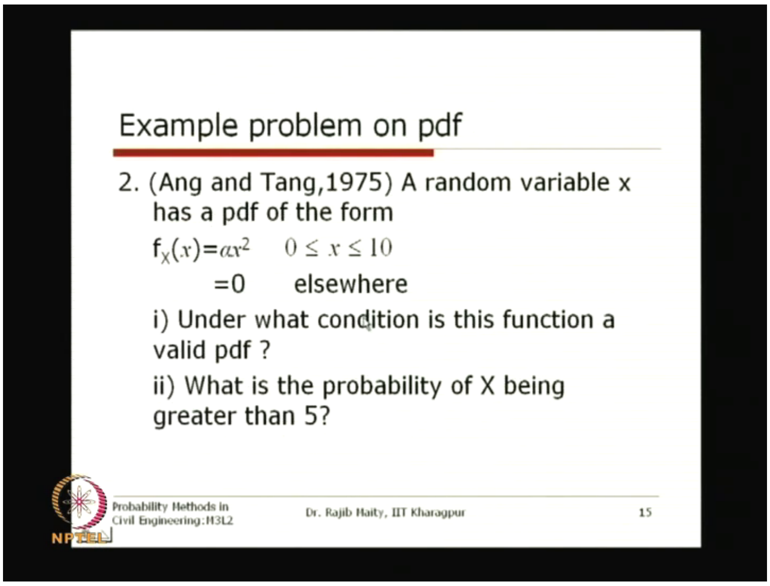
(Refer Slide Time: 36:56)



So second, there are, as we have given some standard example for this pmf probability mass function, which is called the discrete randomvariable.There are some example of some standard pdf as well, which is for the continuous random variable and most popular distribution is a Normal or Gaussian distribution.This Normal Distribution is a continuous probability distribution function with parameters mu and sigma square, this is also, this is known as variance and its probability density function that is pdf is expressed as this one.This is 1 by square root of 2 pi sigma square multiplied by exponential of x minus mu whole square divided by 2 sigma square.

Now, this mu and sigma is known as that parameter of the distribution.Now this, if you change, keeping the basic shape of this probability same, this things are implies different properties of this particular distribution.Before that, what is important, so this is not the complete definition of this pdf, until and unless you say what is its support.So here, the support is minus infinity to plus infinity.So, in absence of this one, basically no function is a valid pdf. So, whenever you are defining some pdf, the support must be must be specified for that function.

(Refer Slide Time: 38:25)



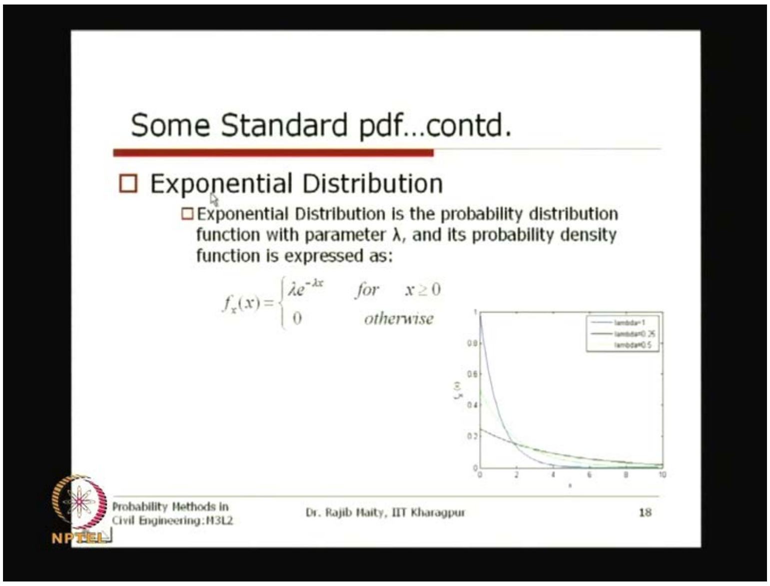
For example, here, so this one, when once you are getting this alpha 3 by 1000, then the pdf is this, is equals to 3 x square by 1000 for x 0 to 10 and equals to 0 elsewhere.So, this support is very important.You know that, if you do not specify this support, then whether this total area below curve is equals to 1 or not, that is that it is cannot be that cannot be tested.So here, similarly, for this normal distribution, the support is minus infinity to plus infinity.So, here some example of this of this normal distribution is shown here.

This is basically, a bell shaped curve and depending of this two parameter, this this can be changed.So, generally this mu is the location parameter. I repeat, this mu is the location parameter, where this distribution, where is the centre of this of this and now I use this word centre very crudely.We will discuss all these things, may be in the next class, but this is the location.

Now you see here, there are three different graphs are shown here.All are Normal Distribution, but, for the different parameter value.So this blue one, the mu is 0, so its location parameter is 0 and the black one is again, the mu equals to 0.So, you can see that this point, both are both are the maximum density is located at this 0,so here.Again, the sigma is is the spread; the variance is the spread above that above that mean.

So, this is 1 and for the second one, it is 1.5. So here, you can see the at the spread, the black one is more and for the blue one is less. For the green one here, the mu is equals to 2.So, you can see that, so this is shifted and the center here again is that 2 and sigma is .75, which is lower than this, the first one, this blue one.So, these are called this mu and sigma is called some parameter of this distribution.This Normal Distribution is symmetric, that you can see it is bell shaped and skewness, so all these things, the skewness,mean, variance these things will be discussed in the next class.So, and again the Normal Distribution also in detail we will be discussingin subsequent classes.What we are just telling here is that, this is over the entire support here. The support is from the minus infinity to plus infinity.One function is defined here and if you do this one, here you cannot do this integration. This is not a closedform integration.The numerical integration has proven that this integration, from this integration, from minus infinity to plus infinity is equals to 1. So, the the area below this curve is equals to 1 and this is known as the Normal or Gaussian distribution.

(Refer Slide Time: 41:30)



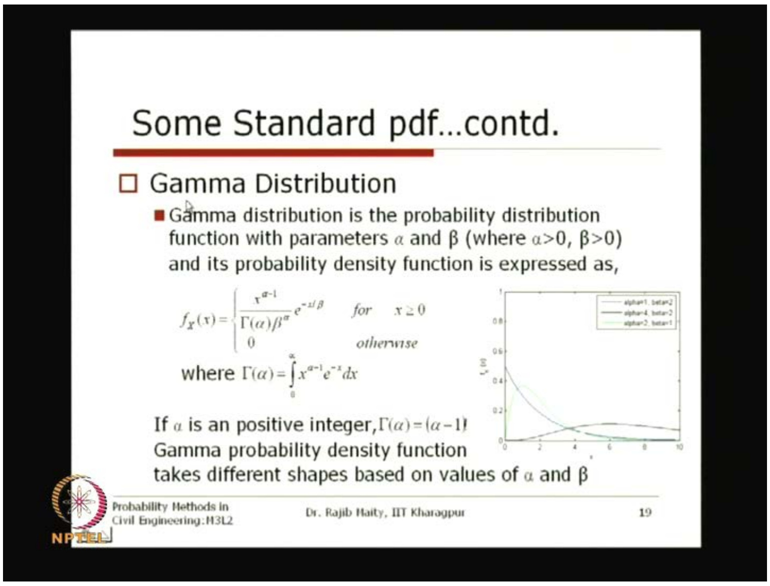
Similarly, another important distribution known as the exponential distribution.This exponential distribution is the probability distribution function with parameter lambda, and its probability density function is expressed as, the f x equals to lambda e power lambda x for x greater than 0.Again, you see, this this support is defined here is greater than equal to 0 and this lambda is known as the parameter of this distribution and it is 0 otherwise.

So, for the entire support, from the minus infinity to plus infinity, this is there. So, this is basically defined in the, defined for the positive x axis.So, these are again some example of this of this exponential distribution for different values of lambda.So, this blue one is showing the lambda is equals to 1.So now, this lambda, this parameter is generally having some relationship with the with the different, as I just discussed the mean and all, this will be discussed in the next class.But, for the time being, this lambda is the parameter for this distribution.So and the difference between this,one difference from this is only one parameter is there as against that normal distribution, where there are two parameters are there.

So, this is single parameter distribution function.This lambda, if we change this lambda, you can see the, if the lambda is equals to 1, this blue curve, the green curve is for the lambda equals to 0.5 and this black one is for lambda equals to 0.25 and these are all defined from this 0 to plus infinity.Now, this integration is very easy.You can just test for these values, if both the integration from 0 to 1, then you will get that the total area below this curve, above this x axis will be equals to 1.

Third one, these are basically, this is whatever the distribution that we are mentioning here, both for this pmf and for the thispdf, this is this is not, that this need be the complete list.Only some examples we are just showing here.You can, some more distribution will be covered in the successive classes as well and here, just we are giving some example, which are generally very important and mostly usedin almost all the field and more importantly, all the fields in the civil engineering.

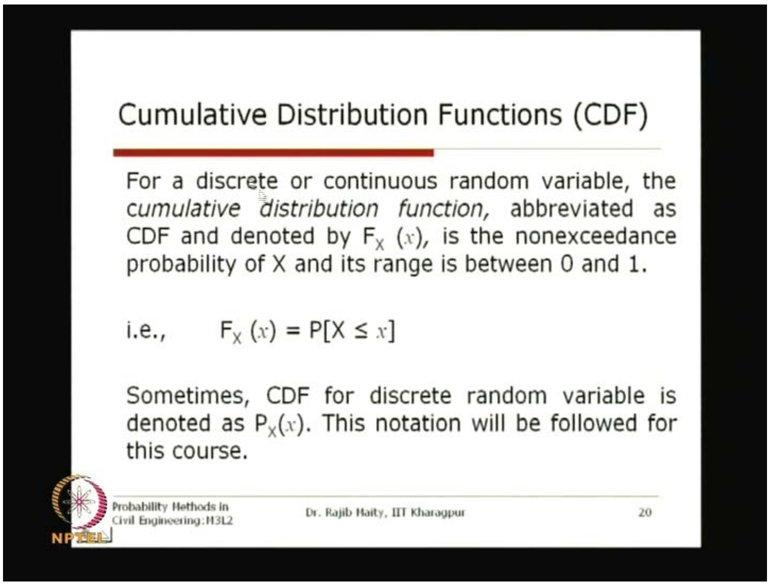
(Refer Slide Time: 43:50)



So, the third example that we are giving is this gamma distribution.Again, this distribution is a two parameter two parameter distribution.The parameters are alpha and beta.So, this is the form of this distribution.This is the one parameter alpha and this is another parameter beta and this is the gamma function.Gamma function again is defined by this integration form and if this alpha is a positive integer, then this form is can be proven.So, this distribution again, this is basically specified for the positive x axis, for the negative side this is 0.this is 0Now, if you see, there is one interesting point here.If you just see that, if you change this alpha to be equals to 1, then, this is nothing but, so alpha equals to 1 gamma alpha, gamma alpha is equals to 0 factorial, which is equals to 1 and alpha equals to 1, so 1 by beta.So, this is x for 0.So, this is 1 by beta e power x minus x by beta.Now, if the 1 by beta is lambda, then this is nothing but, lambda e power e power lambda e power minus lambda x.So, this is again, if I put this alpha equals to 1, this is becoming a exponential distribution.

So, here you can see, if you just change this parameter, then this set changes and the first is the blue one, where the alpha equals to 1 and beta equals to 2.As this alpha equals to 1, this is nothing but the exponential distribution.For the green, itis alpha equals to 4 and beta equals to 2 and for the black one, alpha equals to 2 and beta equals to 1.So, these are again different.gammaThis is the gamma distribution with different parameter values, different combination of the parameters.

(Refer Slide Time: 45:43)



Now,the another important thing in this class that we will discuss is the cumulative distribution function.Now, to see it specifically what we have seen now for the discrete as well as for the continuous distribution, we have seen that, for what is the probability for a specific value in case of the discrete and for what is the density of the distribution,and how it is distributed over the range.Now, this cumulative distribution function; so, for all the other, for the pmf you can get the probability directly, for the pdf you cannot get the probability directly.You have to do the integration over the range to get that one.

Now, CDFis the cumulative distribution function.So, basically we are just going on adding of the probabilities starting from the left extreme.So, the lower extreme of the support to the to the higher extreme of the of the support.So, if you just go on adding up the probabilities, the resulting graph will be the cumulative distribution function.So, for a discrete or for a continuous random variable, the cumulative distribution function abbreviated as CDF, upper case CDF, this D stands for the distribution here and denoted by this f x x.Again, this is the random variable and this is the specific value of the random variable is the nonexceedance probability of the x and its range is between 0 to 1.

So, as I was just discussing that we are just going on accumulating this thing this thing from the lower extreme.From the lower extreme, it is 0 and the upper extreme it will be

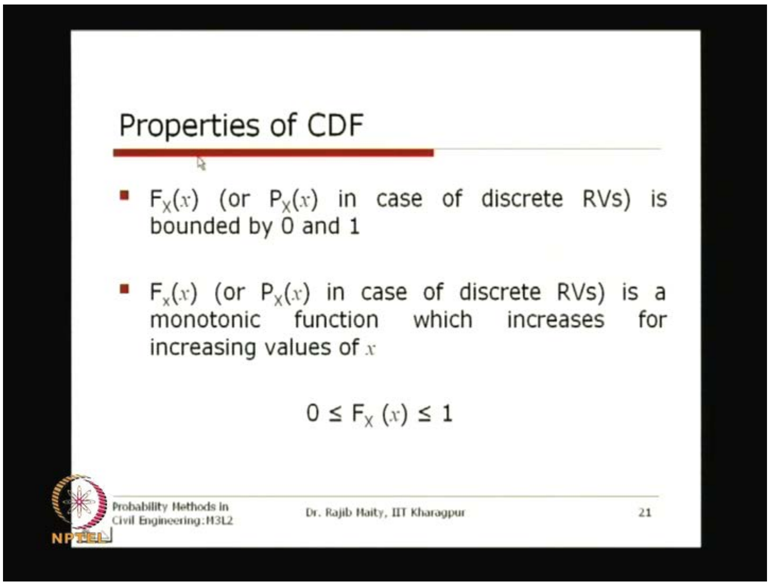
1, obviously. So this f x, as it is stated here is nothing but, the probability for a specific x,probability of the x less than equal to x.So, whatever the whatever the lower value of that specific value of this x, the total probability up to that point is nothing but, this cumulative distribution function.Sometimes, CDFfor the discrete random variable is denoted as p x x.Just this p is, now again the upper case letter and you have seen that pmf, probability mass function, we use this letter as the lower case p.

So, this notation will be followed for this course as well.Now,to again just what we have just telling now, if you just show it here, now this is your now this is your, the probability distribution function that we are doing.Now, what we are trying to say, now to calculate this probability, we have seen that we have to go for this one. Go for this range integration of this range.Instead of that, for this, for the for the CDF, what is what is meant is that we will show for this one.For this specific value, I will calculate this, what is this total area and total area will be put here.So, from here, where the range basically is starting, so this is starting from this 0.Now, we are just going on adding up thesethis values and I will just go on adding, so this value means nothing but, the total area up to this point and in this way we will go on adding. Once, we are reaching here,once we are reaching here we know that total area below.Just now we discussed, the total area below this graph, for a value pdf is equals to 1. So, if you just go on accumulating up to this point, obviously I will reach to the point, where this is equals to 1.

So, obviously that this axis which I have drawn it earlier, for this pdf, need not be the same for this same for this axis. I can just use one more axis system, where it is starts from this 0 and this is ending up to this 1.And, as we are going on adding up these things, obviously this will never will come down.If for some time, if it is, it can go horizontal that it can never come down as this an cumulative function, as the quantities are getting added to this,the to this the earlier value. So, if this is understood, all this concepts for this for this CDFwill be clear.

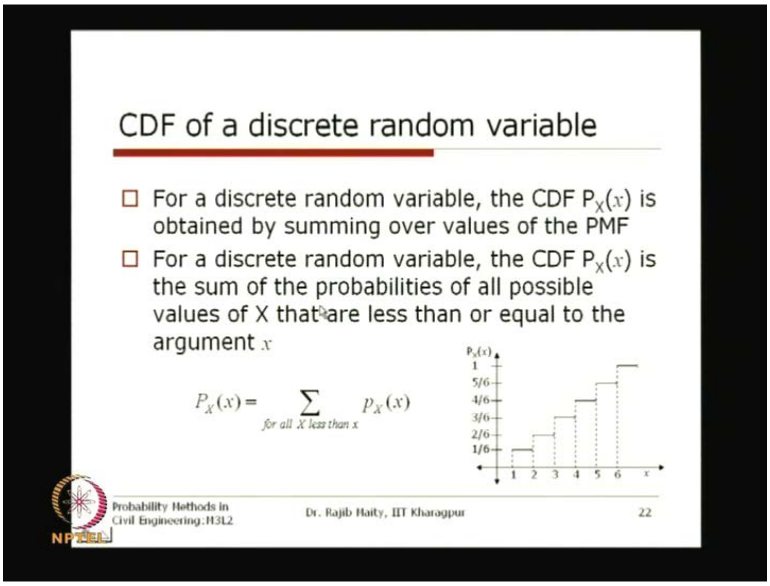
Again, if I get now, if I get this graph, which is CDF, then for a specific value of this random variable, if I want to know what is the probability that x is less, if this one is x, then, what is the probability that that random variable less than equals to that specific value is nothing but, is step forward, we will get it from here, so this will be nothing but, this particular probability what we are getting it here.Now, will go one after another.They are they are the probabilities sometimes.

(Refer Slide Time: 50:45)



So this, now if you see theproperties of this CDF, this f x, that is, which is denoted as the CDFor the p x, in case of the discrete random variable what is just now has told, is bounded by 0 to 0 to 1.So, this is obvious. As you are startingfrom this left let extreme of this support and going up to the right extremes, so it will obviously starts from 0 and it will go up to 1.Secondly, this f x or this p x is a monotonic function, which increases for the increasing values of x.So, this is this is the this is bounded by this 0 to 1.Again, this is monotonic,monotonic function, which increases with this x.This also can be can be clear from this graph, that is, as we are adding up the area, as we are adding up some quantity to the previous value, obviously this function will always increase with increasing values of the x.These two are the properties of the CDF.

(Refer Slide Time: 51:56)



So, now we will just take a take a take a small thing for this one, just to discuss for this district random variable.This is this an important thing, that is, for a discrete random variable, CDFthat is p x is obtained by summing over the values of this pmf. For a discrete random variable, the C D Fp x is the sum of the probabilities of all possible values of x that are less than or equal to the equal to the argument of this x.So, this is this is equals to, for this all values of this x, which is less than x should be added up.If you take the example of the throwing a dice and we know that, for this, all this, there are six equally probable outcomes are there. For all the probabilities,the probabilities 1 by 6, if these are equally probable. Now, if we want to know, what is the CDFfor this one, this will be the starting point of our next lecture and we will see that this is very important to know and where it can touch and where it cannot touch.

This is looks like a step stepfunction.From the next class onwards, we will start a detail description of this one. So in this class, what we have seen is that, we have seen the distribution of the distribution of a particular random variable and this random variable can be discrete, can be continuous.In the next class also, we will see one example.If there are kind of mixed random variable, that also we covered in this that we told in the last class.So here, we will see that how to how to how to handle this issue as a pdf, CDFfor this one, and for the next random variable as well we will see.So, we have first discussed this pmf, probability mass function, which is for the discrete pdf, lower case pdfprobability density function, which is for the continuous and then we have seen how to calculate, how to get the CDFcumulative distribution function from pmf or from the CDF. The concept we have seen and for the discrete one, how to get actual representation of this pmf will be discussed in the next class along with some of the examples taken from the civil engineering problems.Thank you.