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

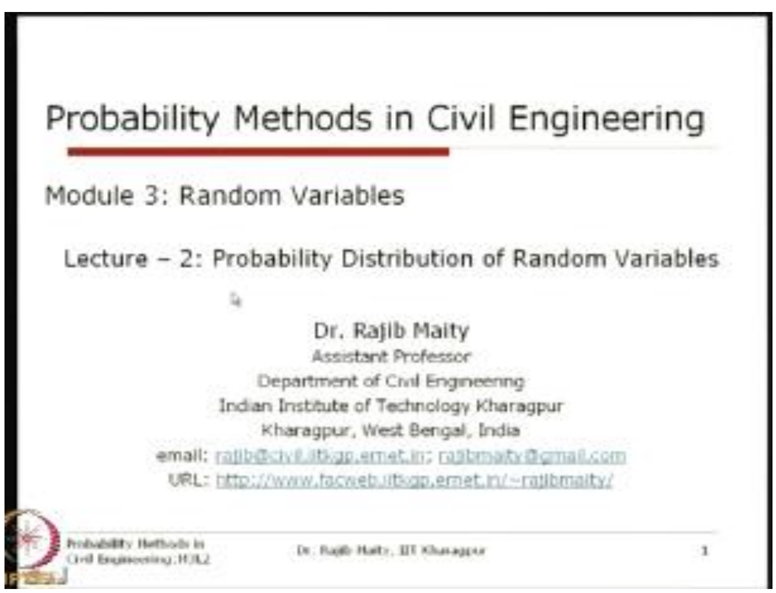
Lecture – 07

Topic

Probability Distribution of Random Variables

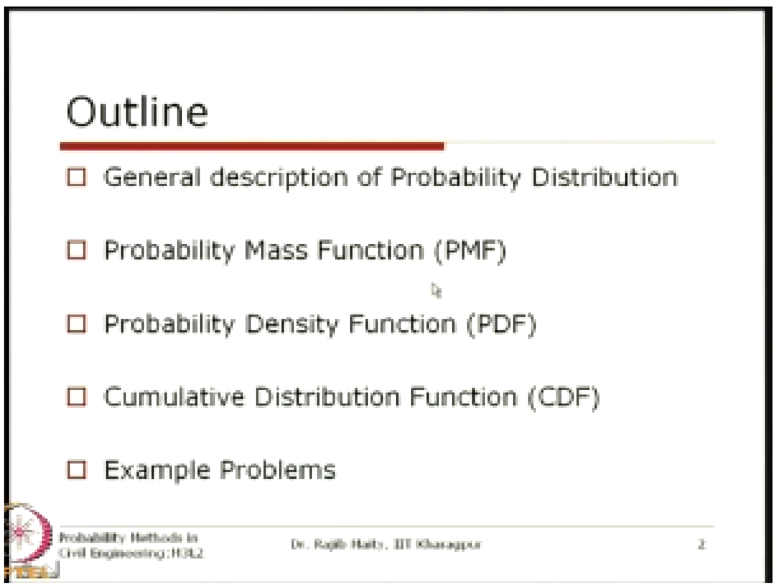
Hello welcome to the second lecture of module three.

(Refer Slide Time: 00:24)



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 theory through its probability distribution. We have to know that over the specific range, over the specified range of one particular random variable, how it is probabilities are distributed. And this is what we will discuss in today's class.

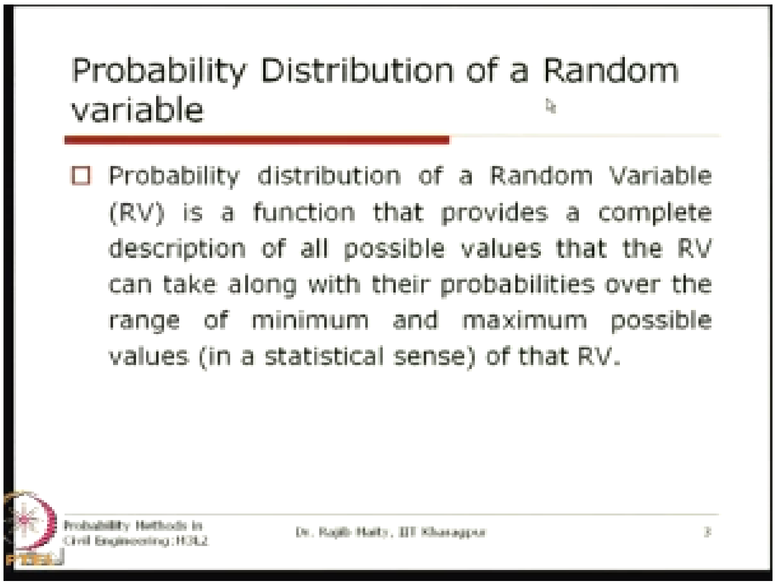
(Refer Slide Time: 00:57)



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 that probability, 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.

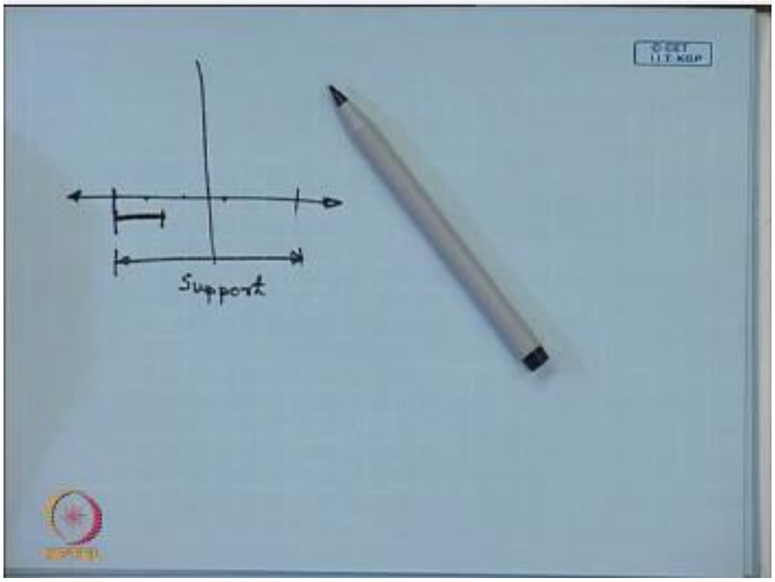
(Refer Slide Time: 01:58)



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 generally having certain range. And over this range, if we have just see it here, over this range of this variables.

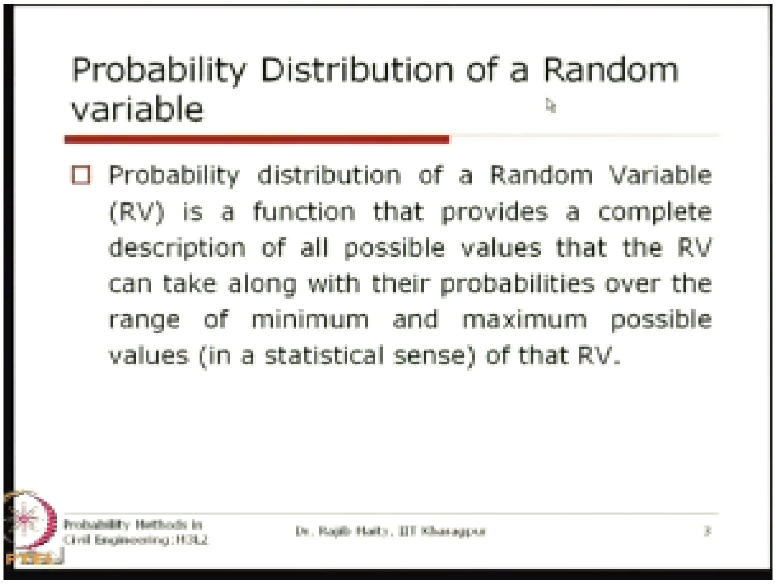
(Refer Slide Time: 03:13)



How 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 your range of this random variable, now thing is that, here how this for each region if it is continuous then for some region, how this probability is distributed over it, for this entire range of this random variable. In the context of this probability distribution, this is known as the support of the random variable.

Support of the random variable and over 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 distributed as a function over this entire support.

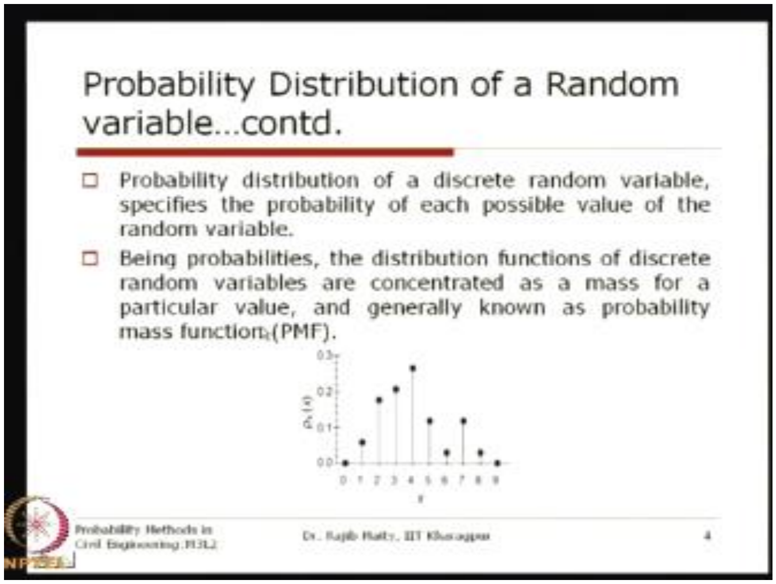
(Refer Slide Time: 04:22)



Thus, so 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 we take, for any random variable, if we just see for one observation, that the maximum of that one or 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. And 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 the distribution of the probability specified for each for all possible values that the random variable can take.

(Refer Slide Time: 05:21)



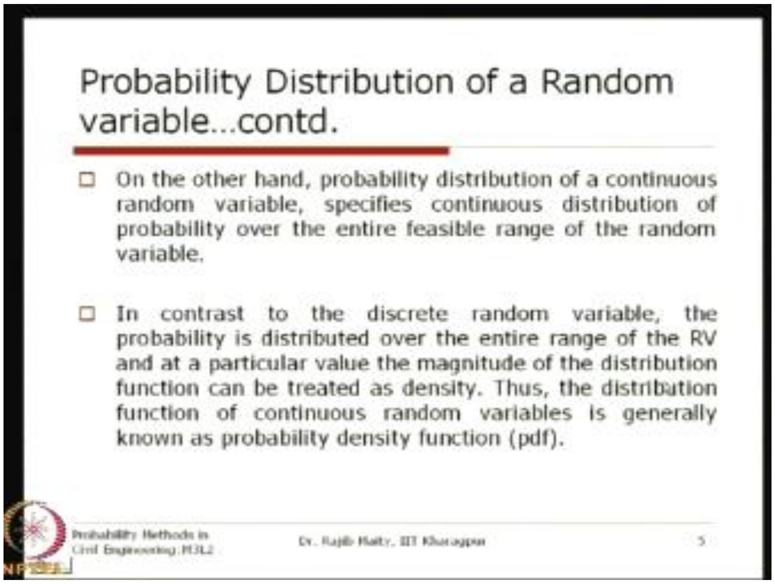
Now as we just mentioning that it can so the two different concept one is for this discrete random variable another one is for the continuous random variable as we discussed in the last class that discrete random variable here means that it can take some specific value over the range of this random variable it cannot take all possible values even though traditionally or in most of the cases this discrete random variable takes the integer value but, that is only the concept it can take any specific value 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 enter support of that of the distribution or entire 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 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 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 is not specified by this by this distribution okay 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 further 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 on the other hand as for.

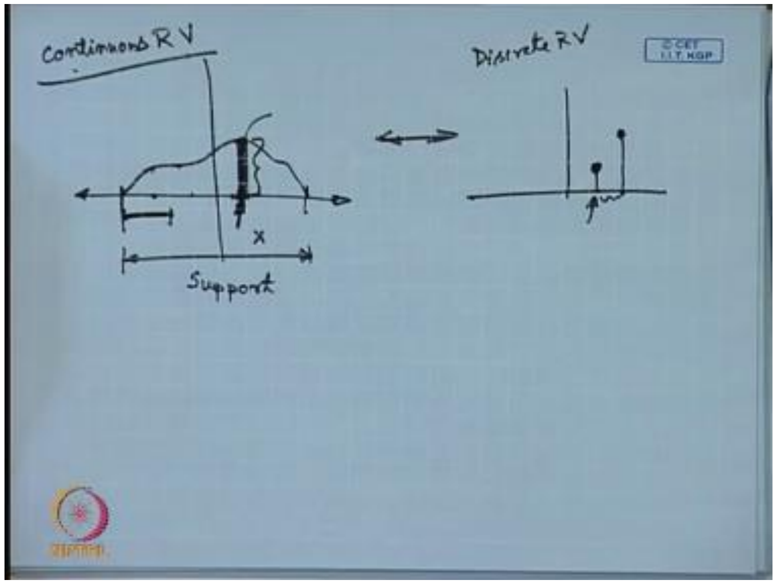
(Refer Slide Time: 08:35)



The for the continuous random variable this is not the case so this can be this is specified over the over a region so on the other hand the probability distribution of a continuous random variable specified 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 variables is generally known as probability density function and abbreviated as lower case of PDF. So this point that is the its concept of the density as I will just explain it here as I was telling.

(Refer Slide Time: 09:51)



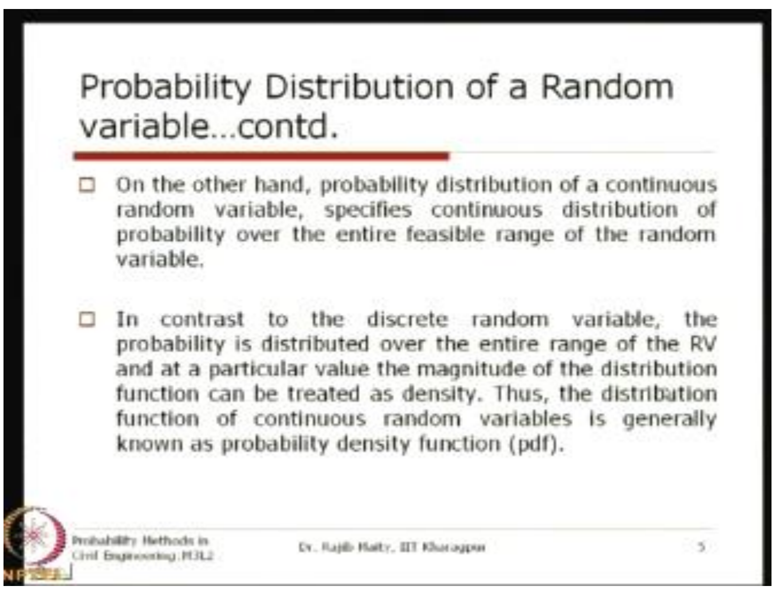
That if this is the entire support of the so now what I am discussing this is in case of the continuous random variable okay so for this continuous random variable if I say that okay fine this distribution is specified distribution is shown by this so for what it first of all what is what is showing that for some region 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 implies is this implies the probability now if I just 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 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 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 normal physical science if you multiply that density with it is mass then you will get it over.

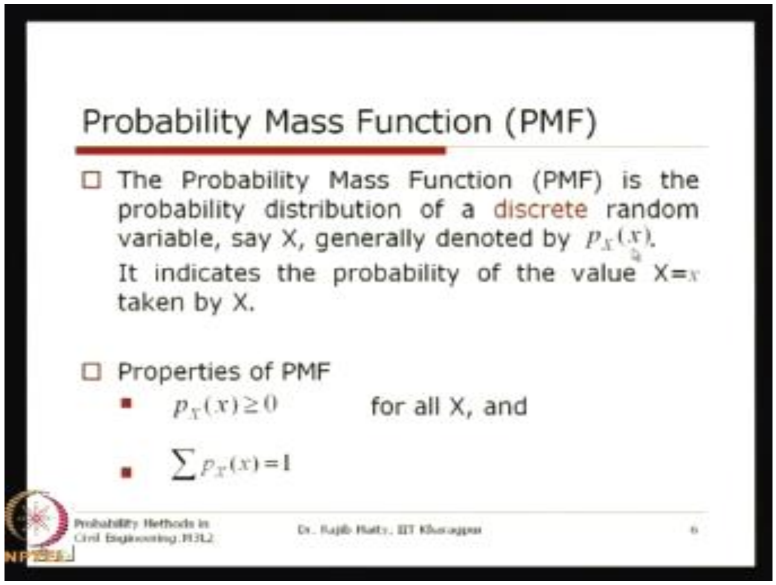
So similarly here this is nothing but the density if you multiply over a certain range then we will get the area and that area nothing but your probability so that is why for this distribution for the continuous random variable this distribution is the probability density function so this is why the word density comes here so once again if we just read it.

(Refer Slide Time: 12:29)



In contrast to the discrete random variable the probability is distributed over the entire 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 will be more clear when we are talking about that one of the axiom that we have seen in the 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: 13:14)



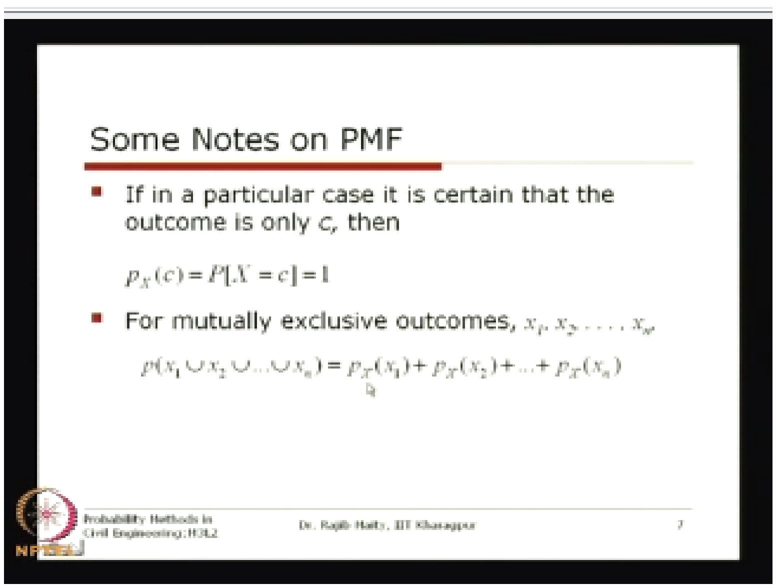
So next we will first start with this probability mass PMF which is for this we just now have seen this is for the discrete random variable. So whenever we mention that PMF this is for the discrete random variable the probability mass function PMF is the probability distribution of a discrete random variable say X generally this is denoted by P, X, X. So here this notation is important this p is the lower case this 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: 15:26)



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.

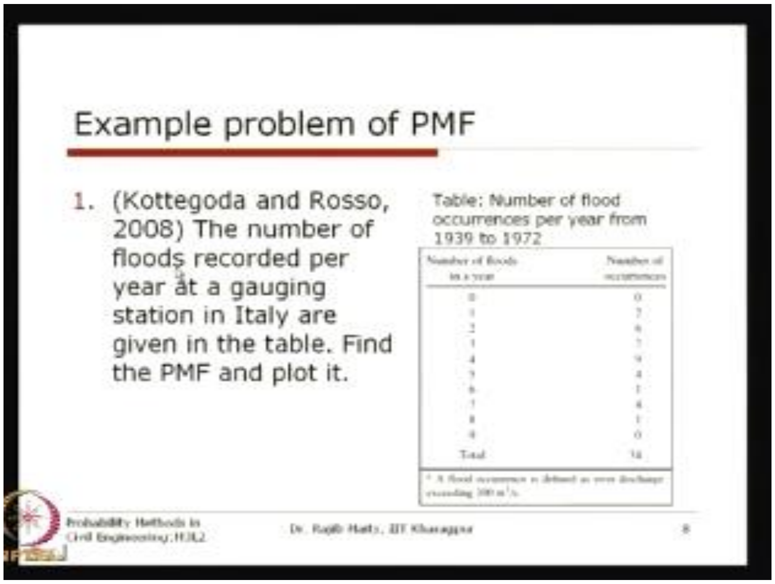
This is the only outcome that is that is feasible and this is certain outcome so this will obviously come up, so this is equals so this should entirely be equal to 1, to satisfy all this properties of this PMF. On the other hand if there are some mutually exclusive outcomes for one random variable that is x1, x2….xn. Now if we just say that these this 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 value that the random variable contains should be equal to the summation of their individual probabilities. So this is obvious in case of a throwing a dice and there are six possible outcomes and if you say that all the outcomes are equally feasible equally 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: 17:45)

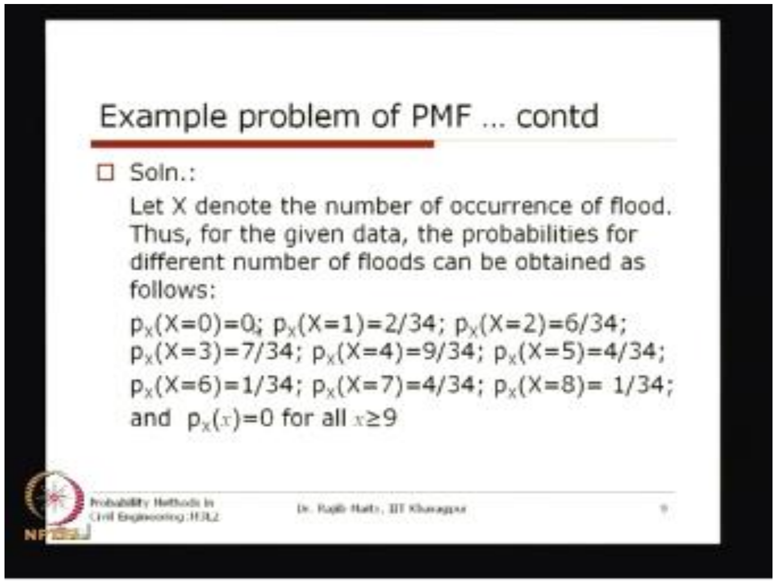


So one small example will take on this on this probability mass function this is taken from the Kottegoda and Rosso book. Kottegod and Rosso book, The number of floods recorded per year at a gauging station in Italy are given in this table. Find the probability mass function and plot it. So now here for this 1939 to 1972 there are so many number of floods are noted here so 0 floods so this table is this.

So 0 flood has occurred in 0 years. So and 1flood has occurred in 2 years, 2 floods has 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 utmost 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 we should think.

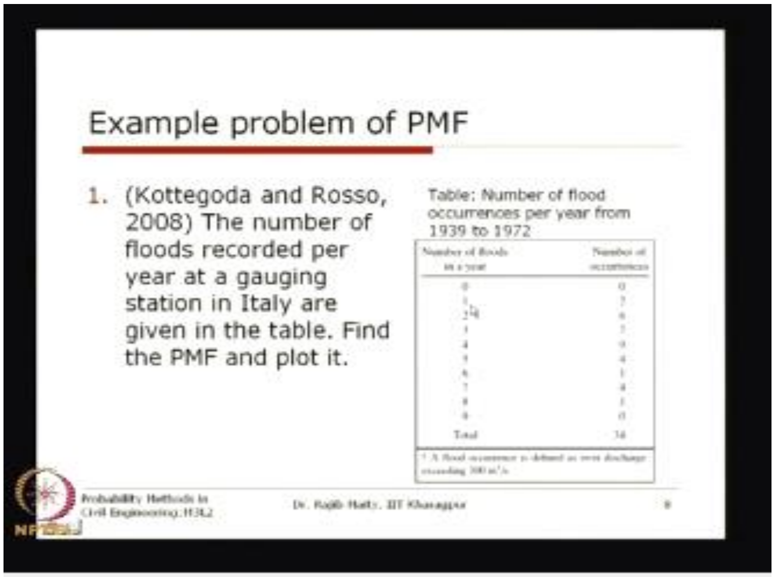
That what is the random variable that we are talking about? So this is one of the even though this is sometime for this kind of problem this is quite obvious but this is important to know that what is the random variable random variable here. So occurring the flood is not the random variable I repeat,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

(Refer Slide Time: 19:24)



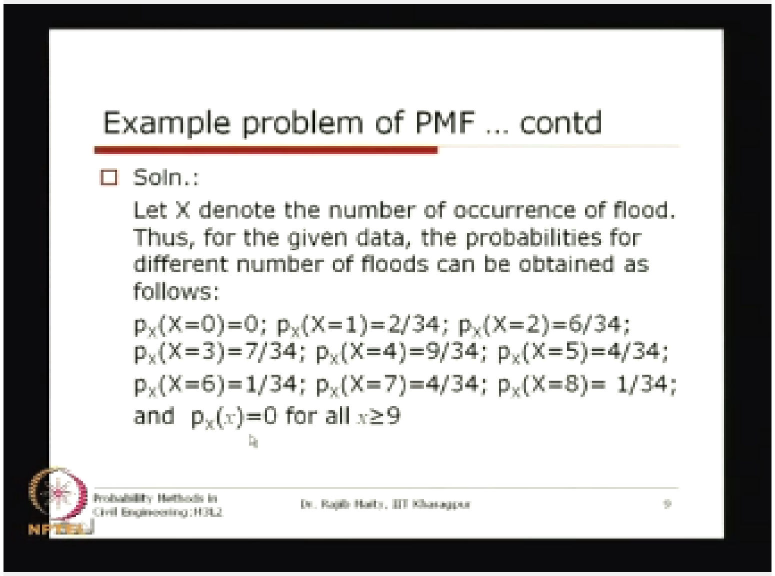
So while solving this problem first of all what we will do, we will first define that what is the random variable that we are mentioning. So let x that is the random variable denote the number of occurrence of flood, thus for the given data the probabilities of the different number of the floods can be obtained as follows. So p(x) = 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.

(Refer Slide Time: 19:58)



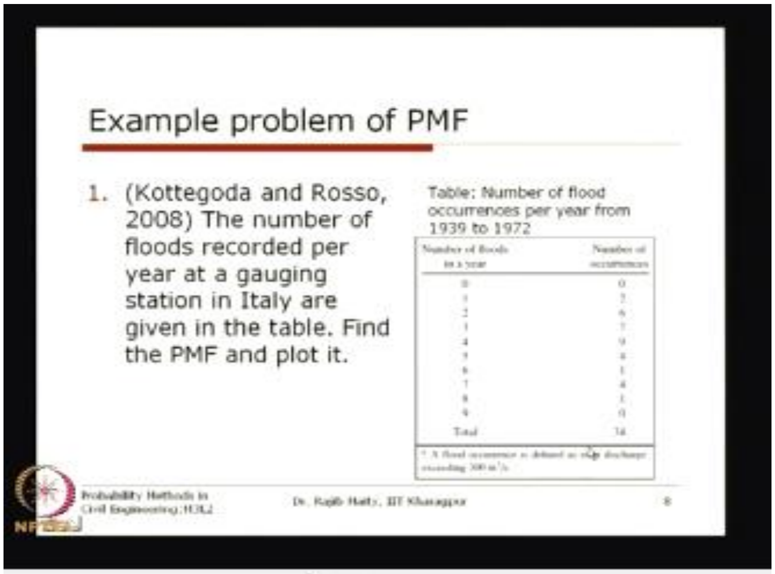
Because from this table we see there is no such a year where the number of flood is 0.Similarly, if we want to know the what is the probability that x equals to 1 that means from this sample data obviously so x = 1 so that means there are two such occurrences out of 34 years so 2/34 should be the probability for that specific value of the random variable, that is x = 1.

(Refer Slide Time: 20:23)



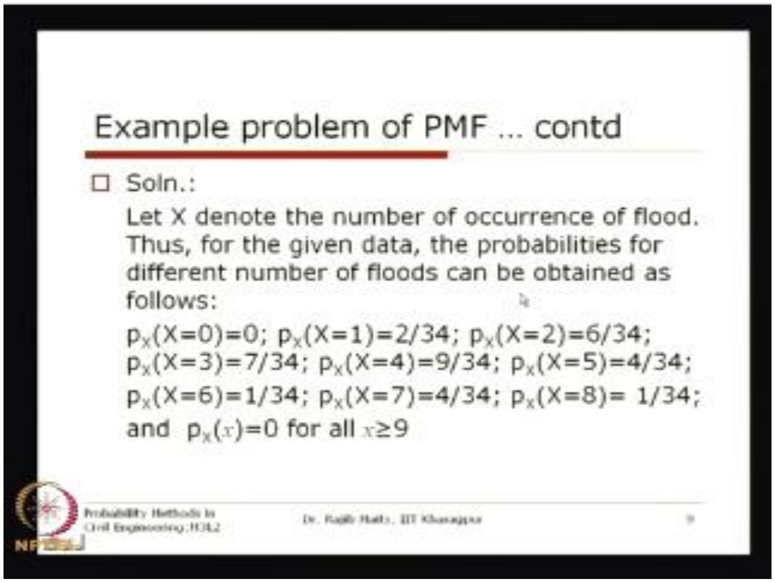
Similarly so this is 1 the probability of x = 1= 2/34.Similarly for x 2, 6/34, x = 3, 7/34 and in so on we are going and getting all this probability values.

(Refer Slide Time: 20:43)



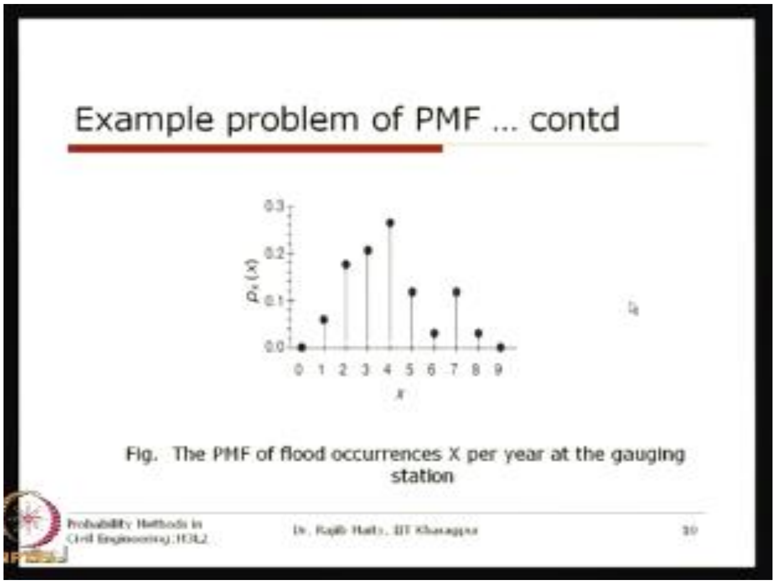
Now, similarly we are assuming one thing that as so number of flood 9 floods in a in a year this is occurring 0 and all we just got the summation this one. So there is no other occur numbers of flood if I just take 10, 11 there are also number of occurrence in a year for 10 floods in a year 0,11 floods in a year is 0.

(Refer Slide Time: 21:03)



So to complete the definition of this PMF x = 0 for all x > 9. Now, simply we have to plot this one as a 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.

(Refer Slide Time: 21:29)

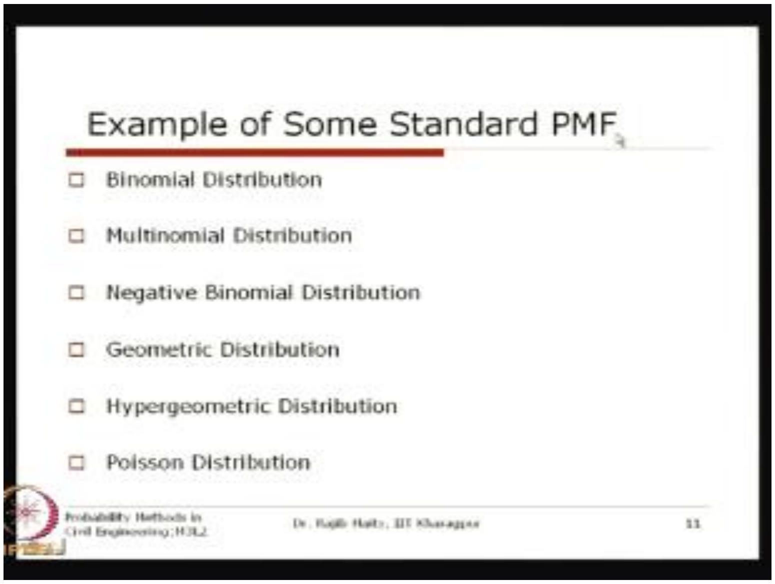


So this is the plot So for the 0 the probability is 0 and obviously for 9, probability is 0, for all higher values is also 0.For 1, these are the probabilities whatever the value we got here so we are getting this masses. Now, obviously do not confuse that this what is the meaning of this line, basically this line this solid line has no meaning just for as a geometric reference that this point refers to this number1.

Otherwise a simple a single dot at this point should be sufficient to display the probability mass function then now we can form this probability mass function we got the probability mass function to the data that we have Several things can be answered from this one but 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 and this will again see while we are discussing the cumulative distribution function.

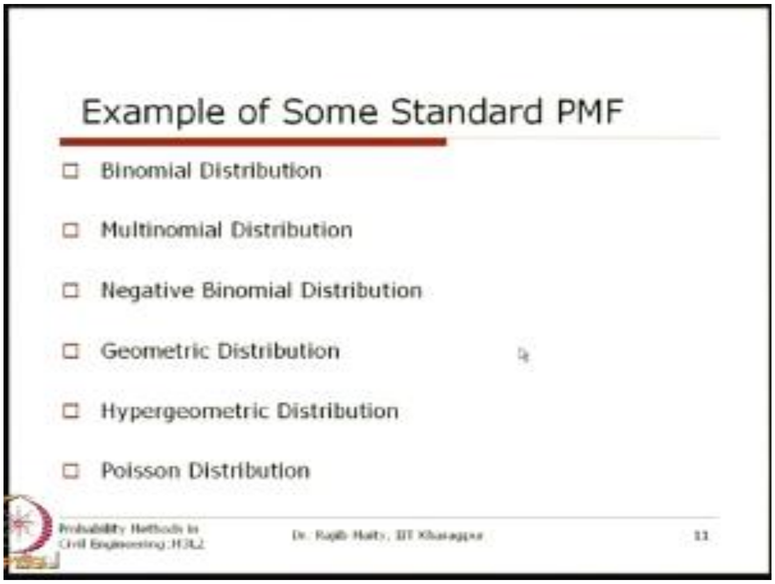
(Refer Slide Time: 22:57)



So this is for the discrete and there are some standard examples 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 bernoullitrial where there are two possible outcomes.

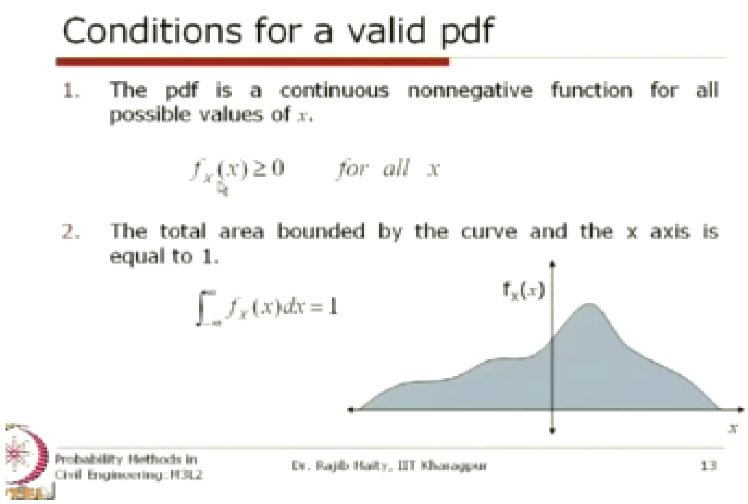
And one we just we just tell it a successes and the success for each trail is predefined which is known so now the number of number success out of n successive such trails, that is the number, so that number is a random variable and that random variable follow this binomial distribution similarly if there are more than two more than two outcomes then for how many success we are getting in a set of say 1 to k and all this success rates are known then the vector that is that x 1 x 2 x 3up to x k.

(Refer Slide Time: 24:00)



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 hyper geometric 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.

(Refer Slide Time: 24:29)



Random variable 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 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 abbreviated as lower case of pdf.

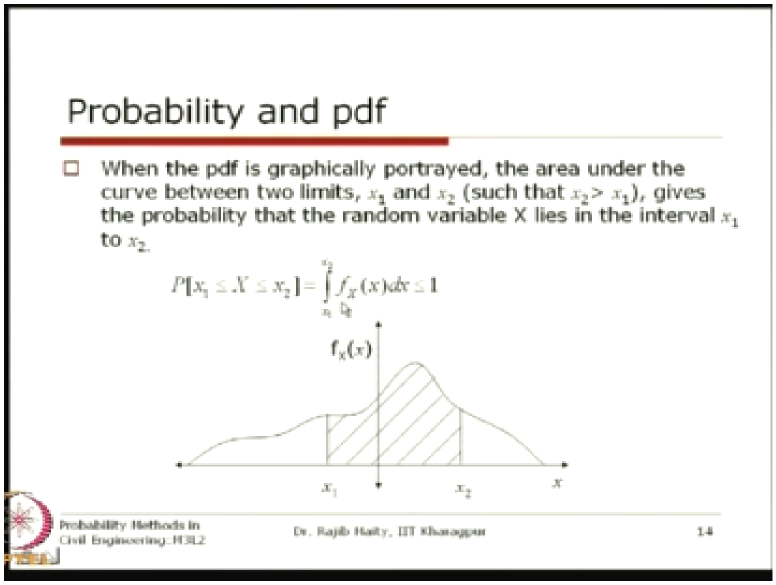
Just to differentiated from this cumulative distribution function where the d stands for the distribution here this d stands for the density so this pdf is the probability distribution of continuous random variable generally it is denoted by this f x x which we discussed last time also, that this is 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 followed to make a particular function to be a probability density function and those conditions are this so there are two condition conditions for a valid pdf the pdf 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 on 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 -∞ plus ∞ 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 rest of 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.

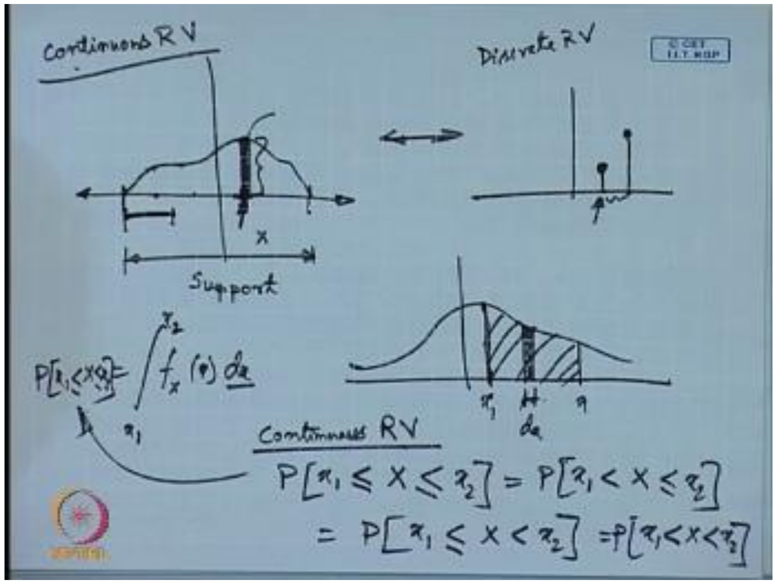
(Refer Slide Time: 28:05)



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 x1 and x 2 such that say x 2is greater than x1 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 x2 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 to1 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 so the integration here.

(Refer Slide Time: 29:34)

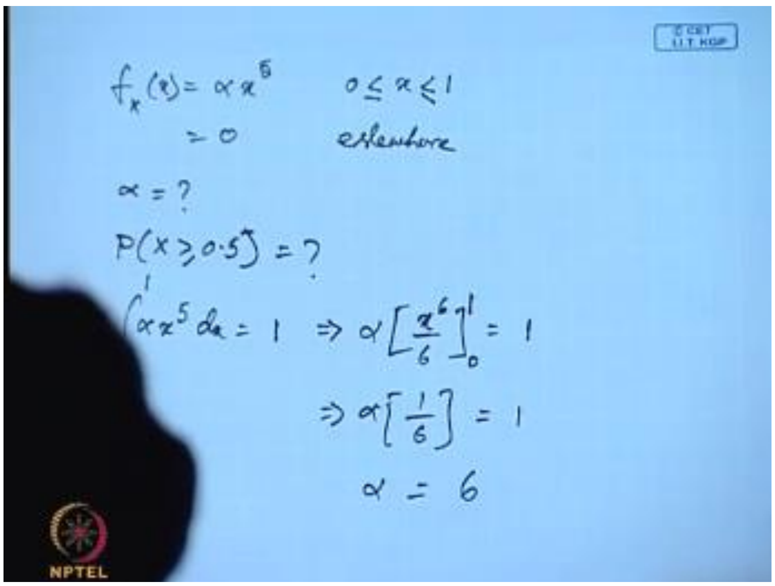


Is so integration what we have seen here is that f x x d x from x1 to x 2. now what we have just discussing here that this is this is the density here basically what we are doing here for this graph so we are taking a small so d x is your small strip here dx is the small strip here so this is your d x and the this one 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 x1 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 x2.

And that is nothing but which is gives you the gives the probability 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.

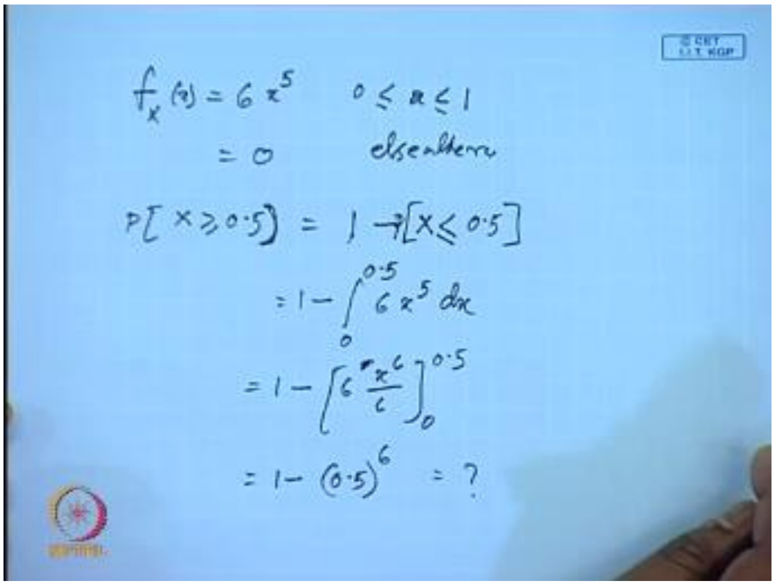
(Refer Slide Time: 32:27)



Now we will take one small example mathematical example rather to discuss about this pdf how to satisfy their properties suppose that the function $f_x \; x = \alpha x^5$ and which is defined over the zone, for this x value from 0 to1 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.

So if you want to know this first one that is the , what is the value of this α 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 α this x^6 / 6, which is equals to your1.This is α so, it is 0 to (01 /6) –(0 = 1), where the alpha is equals to6. This value we got so now, what we will see that.

(Refer Slide Time: 34:09)

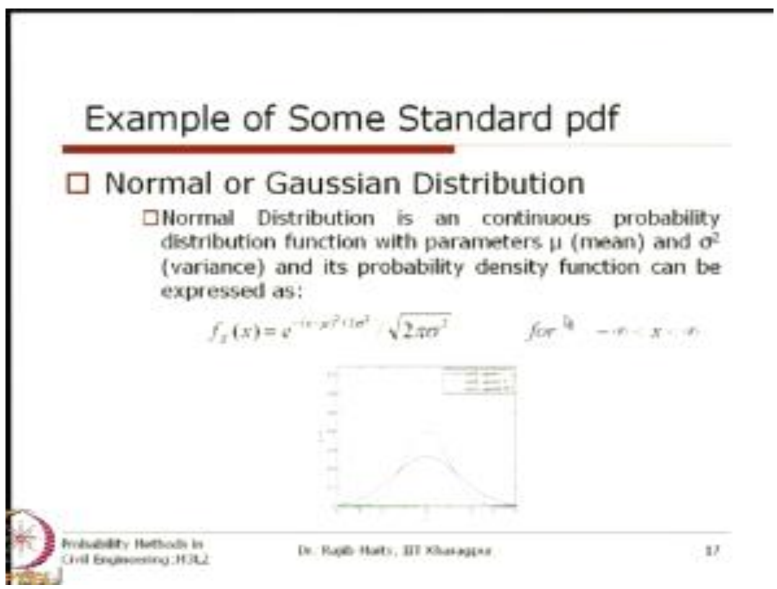


This so what we got that f x of this x = $6x^5$ over the range this x 1 and equals to 0 elsewhere so now the second thing is that what is the probability that x is greater than equal to 0.5.This we can express that you know that 1 minus probability of x less than or less than equals t, you know that 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^5 d x, which is 1- 6 / 6.We can just write to here, 0 to 0.5 so it is equals to 1 minus, say 0.5^6. 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 to1, because this is the range 0.5 to 1we 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 1so1 minus the rest of this part, from this means here 0to 0.5.What is the area that we have deducted to get this probability basically this relates its link to the to the CDF because we will just see in a minute that what is a CDF. So from the CDF, 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 CDF that we are going to discuss in a minute. But so far as this particular problem is concerned, we can also calculate the integration from0.5 to 1 to get this particular probability answer.

(Refer Slide Time: 36:51)

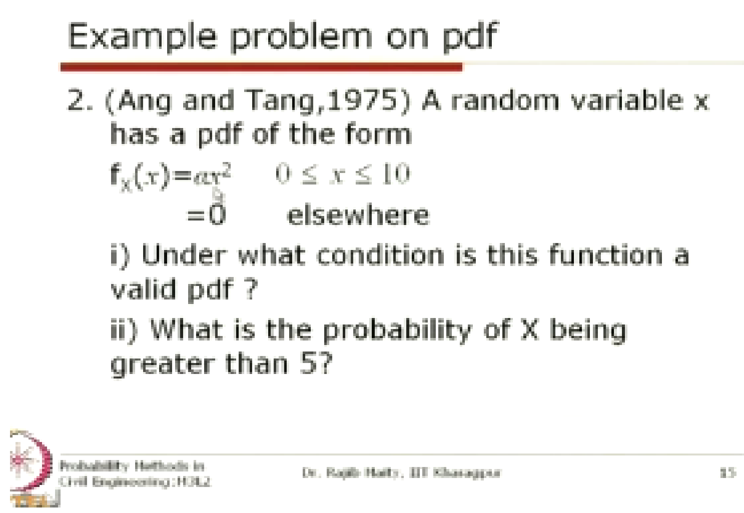


So second there are as we have given some standard example for this pmf probability mass function, which is called the discrete random variable. There are some example of some standard pdf as well which is for the continuous random variable and most popular distribution is a Normal or Gaussian distribution. 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 $1/\sqrt{2\pi\sigma^2}$ x 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 its support is.

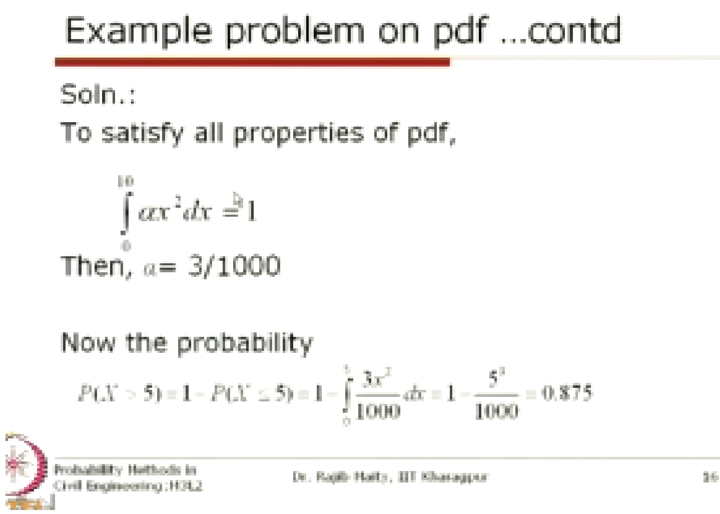
So here the support is minus infinity to plus infinity. So in absence of this one basically no function is a valid pdf. So whenever you are defining some pdf the support must be must be specified for that function.

(Refer Slide Time: 38:25)

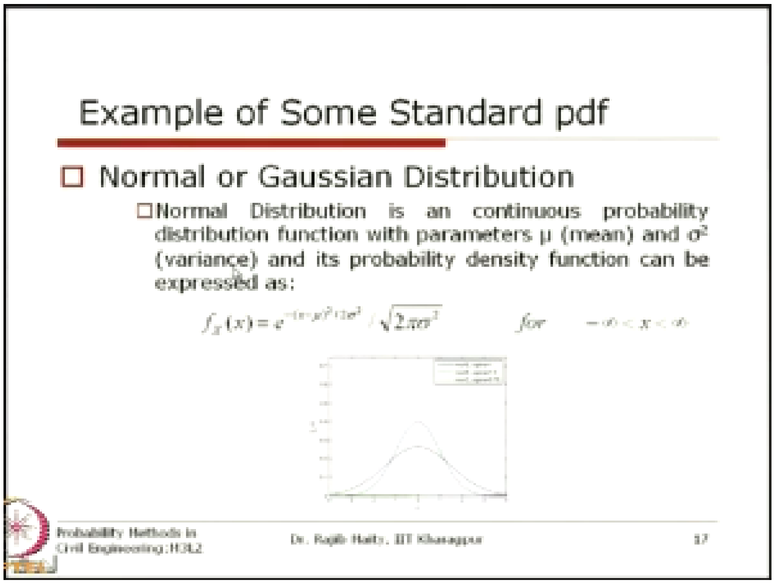


For example here so this one when once you are getting this α 3 / 1000, then the pdf is this, is equals to 3 x square by1000 for x 0 to 10 and equals to 0 elsewhere. So this support is very important you know that, if you do not specify this support then whether this total area below curve is equals to 1or not, that is that it is cannot be that cannot be tested.

(Refer Slide Time: 38:50)



(Refer Slide Time: 38:51)



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 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.

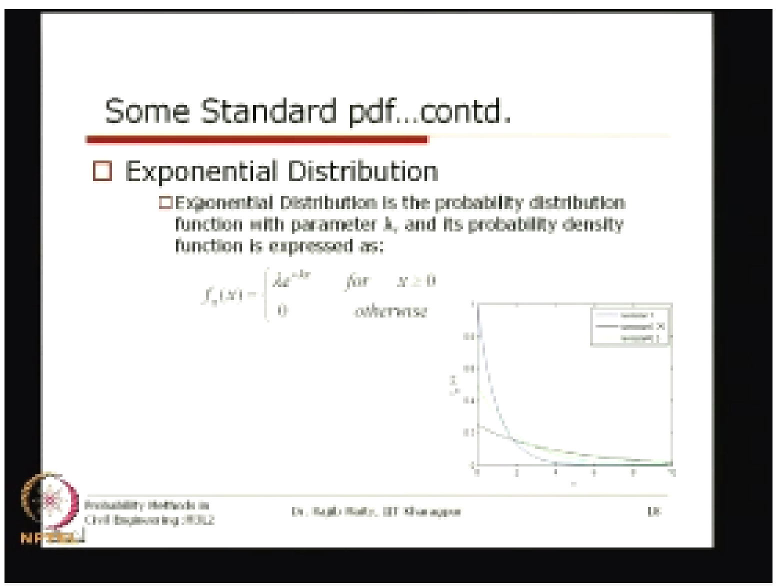
But this is the location now you see here there are three different graphs are shown here all are Normal Distribution but for the different parameter value. So this blue one the mu is 0 so its location parameter is 0 and the black one is again the mu equals to 0.So you can see that this point, both are both are the maximum density is located at this 0so here.

Again the σ is the spread the variance is the spread above that above that mean so this is 1 and for the second one it is1.5. 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 that2 and σ is .75, which is lower than this the first one, this blue one. So these are called this mu and σ is called some parameter of this distribution.

This Normal Distribution is symmetric, that you can see it is bell shaped and skew ness, so all these things the skew ness, mean, variance these things will be discussed in the next class. So and again the Normal Distribution also in detail we will be discussing in 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 closed form integration. The numerical integration has proven that this integration, from this integration, from minus infinity to plus infinity is equals to 1. So, the area below this curve is equals to 1 and this is known as the Normal or Gaussian distribution.

(Refer Slide Time: 41:24)



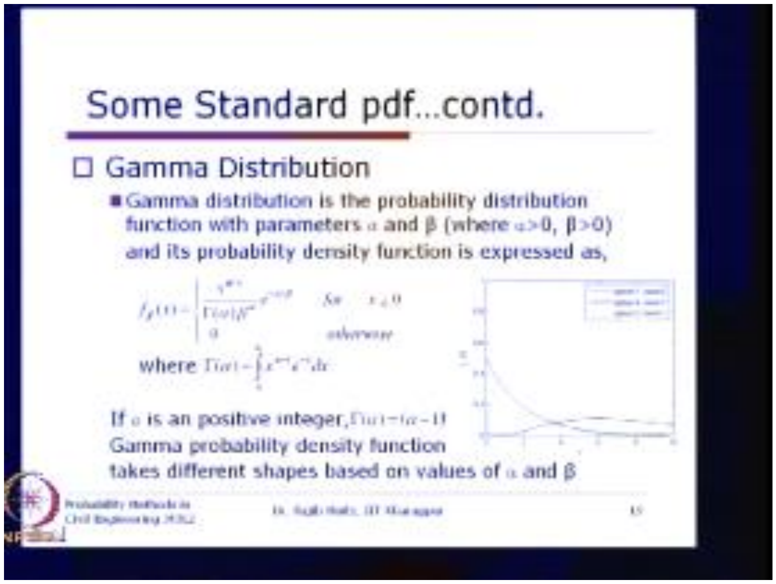
Similarly another important distribution known as the exponential distribution. This exponential distribution is the probability distribution function with parameter λ, and its probability density function is expressed as, the fx equals to λ e power λ x for x greater than 0.Again, you see, this this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

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 **λ**. So, this blue one is showing the **λ** is equals to 1.Sonow, this **λ**, 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 **λ** 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 **λ**, if we change this **λ** ,you can see the, if the **λ** is equals to1, this blue curve, the green curve is for the **λ** equals to 0.5 and this black one is for **λ** 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 this pdf, 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 used in almost all the field and more importantly, all the fields in the civil engineering.

(Refer Slide Time: 43:46)

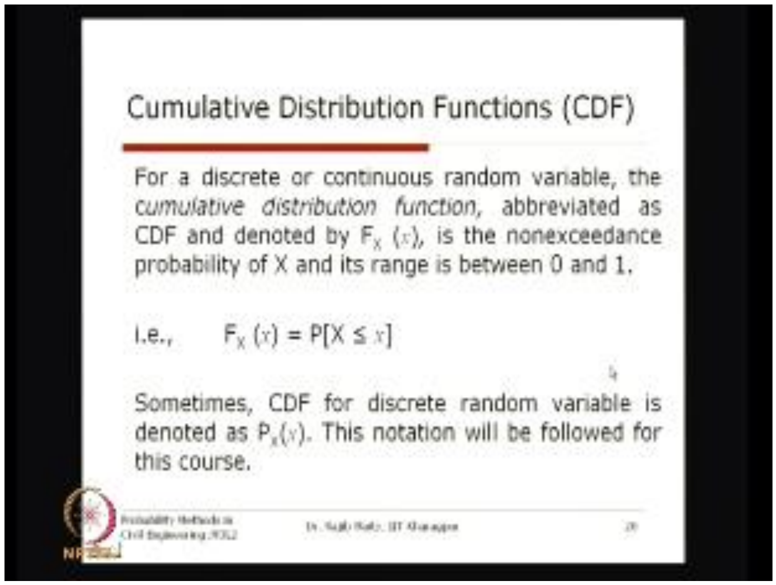


So, the third example that we are giving is this γ distribution. Again, this distribution is a two parameter two parameter distribution. The parameters are α and β. So, this is the form of this distribution. This is the one parameter α and this is another parameter β and this is the **γ** function. Γ function again is defined by this integration form and if this α 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 α to be equals to 1, then, this is nothing but, so α equals to 1 **γ** α, **γ** α is equals to0 factorial, which is equals to 1 and α equals to 1, so 1 by β. So, this is x for0. So, this is 1 by β e power x minus x by β Now, if the 1 by β is λ, then this is nothing but, **λ** e power **λ** e power minus **λ** x.

So, this is again, if I put this, α 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 α equals to1 and β equals to 2.As this α equals to 1, this is nothing but the exponential distribution. For the green, it is α, equals to 4 and β equals to 2 and for the black one, α equals to 2 and β equals to1.So, these are again different. Γ This is the γ distribution with different parameter values, different combination of the parameters.

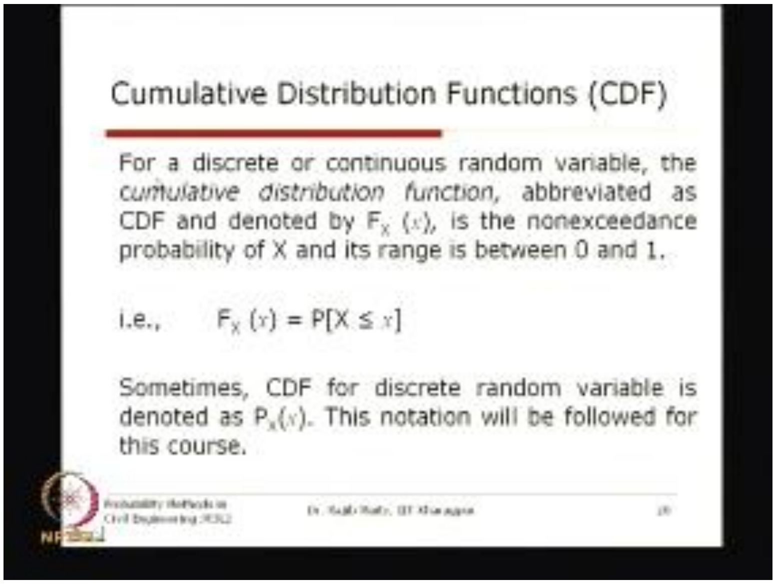
(Refer Slide Time: 45:38)



Now, the 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, CDF is 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.

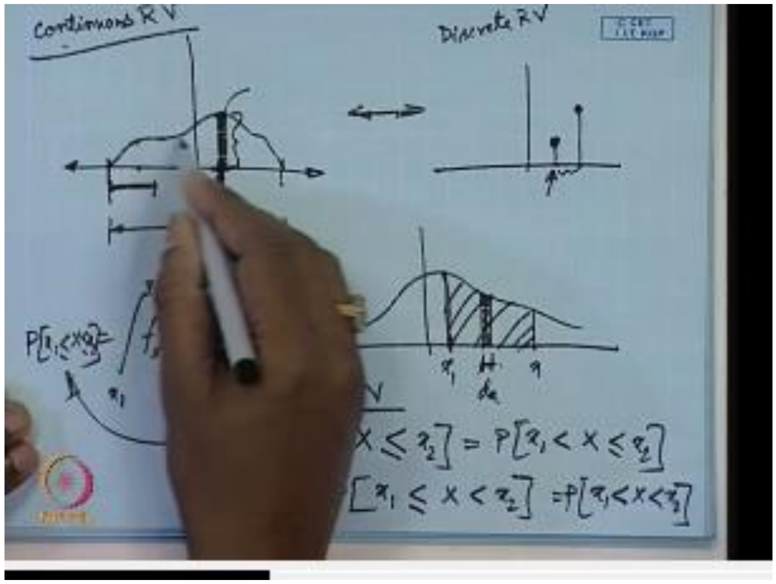
(Refer Slide Time: 46:40)



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_{xx} .Again, this is the random variable and this is the specific value of the random variable is the non-exceed an 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 fx, as it is stated here is nothing but, the probability for a specific ,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, CDF for the discrete random variable is denoted as Pxx. Just this p is, now again the uppercase letter and you have seen that P_{mf} , probability mass function, we use this letter as the lowercase 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

(Refer Slide Time: 48:10)

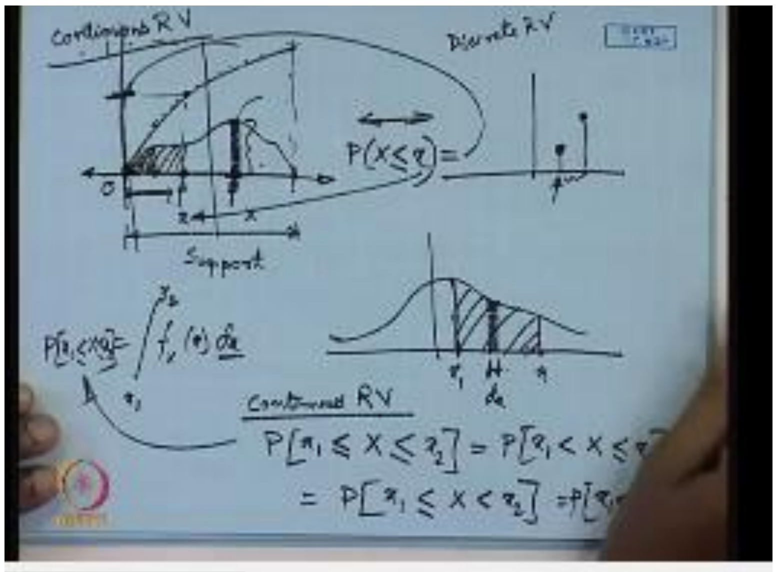


Now this is your, the probability distribution function that we are doing. Now, what we are trying to say, now to calculate this probability, we have seen that we have to go for this one. Go for this range integration of this range. Instead of that, for this, for the , for the CDF, what is what is meant 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's, from here, where the range basically is starting, so this is starting from this 0.

Now, we are just going on adding up these this 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.

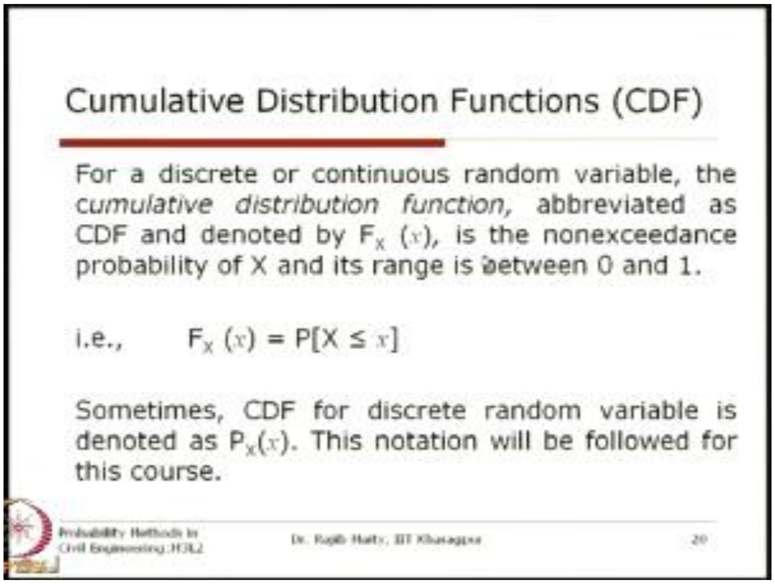
I can just use one more axis system, where it is starts from this 0 and this is ending up to this 1.And, as we are going on adding up these things, obviously this will never will come down. If for some time, if it is, it can go horizontal that it can never come down as this an cumulative function, as the quantities are getting added to this, the to this the earlier value. So, if this is understood, all this concepts for this.

(Refer Slide Time: 50:03)



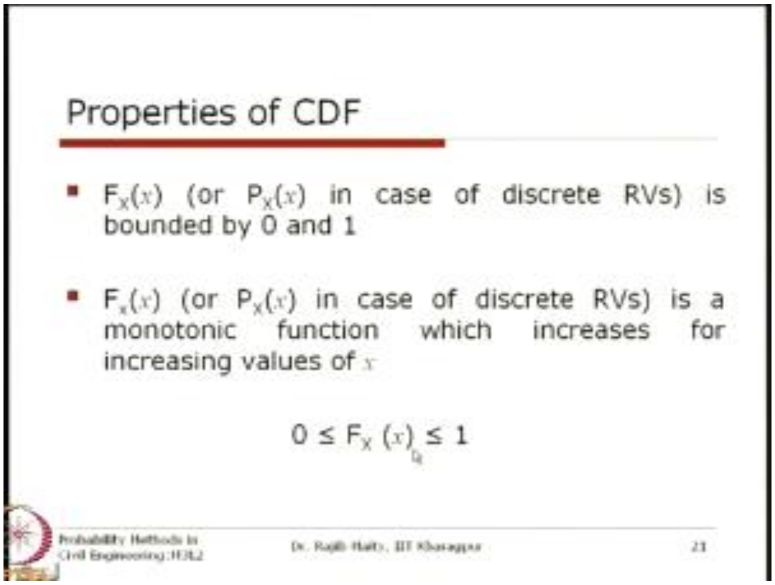
If this is understood all this concepts for this CDF will be clear. Again if I get now if I get this graph which is CDF then for a specific value of this random variable if I want to know what is the probability that x is less so if this one is x, then what is the probability that random variable less 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.

(Refer Slide Time: 50.35)



So now we will go one after another they are the probabilities sometimes.

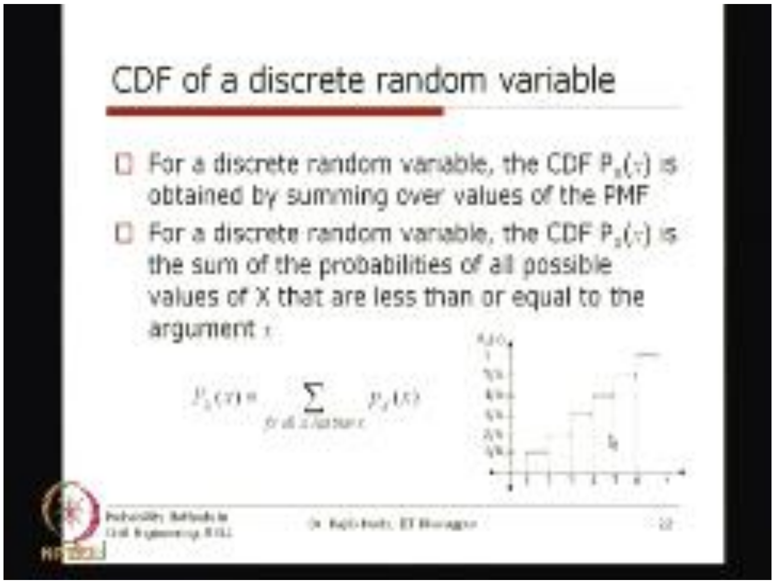
(Refer Slide Time: 50:41)



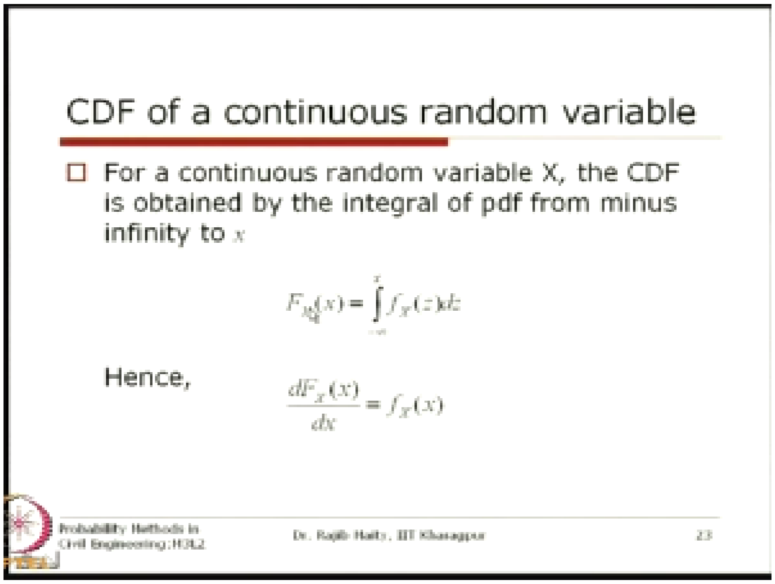
So this, now if you see the properties of this CDF this fx that is which is denoted as the CDF or the p x in case of the discrete random variable what is just now has told is bounded by 0 to 1. So this is obvious as we are starting from this left extreme of this support and going up to the right extremes so it will obviously starts from 0 and it will go up to 1. Secondly this fx or this px is a monotonic function, which increases for the increasing values of x.

So this is bounded by this 0 to 1 again this is monotonic, a monotonic function which increases with this x this also 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. So these two are the properties of the CDF.

(Refer Slide Time: 51:53)

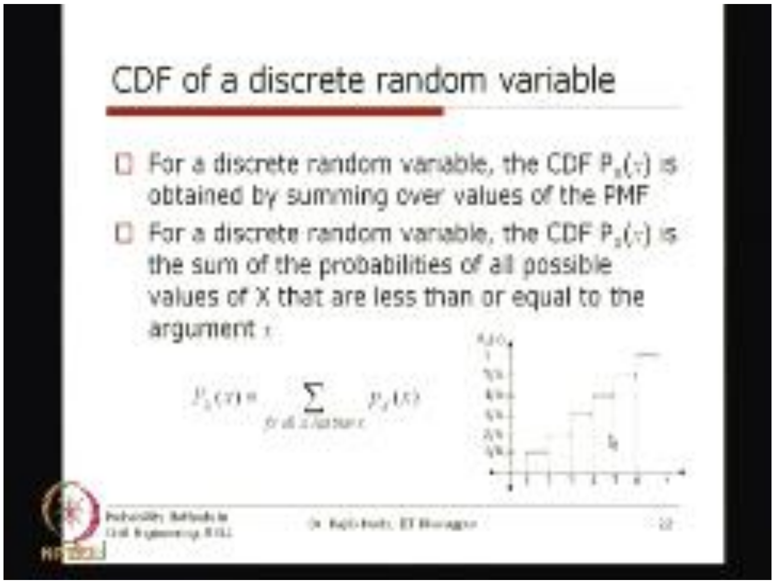


(Refer Slide Time: 51:58)



So now we will just take a small thing.

(Refer Slide Time: 52:00)



For this one just to discuss for this discrete random variable this is this an important thing that is, for a discrete random variable CDF that is px is obtained by summing over the values of this PMF for a discrete random variable the CDF px is the sum of the probabilities of all possible values of x that are less than or 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 we 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 CDF for 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 looks like a step function and 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 particular random variable and this random variable can be discrete, can be continuous, and 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 handle this issue as a pdf, CDF for 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 pdf probability density function which is for the continuous and then we have seen how to calculate, how to get the CDF cumulative 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.

Probability Methods in Civil Engineering

End of Lecture 07

Next: "CDF and Descriptors of Random Variables" In Lec 08

NPTEL Video Recording Team
NPTEL Web Editing Team
Technical Superintendents
Computer Technicians

A IIT Kharagpur Production
www.nptel.iitm.ac.in