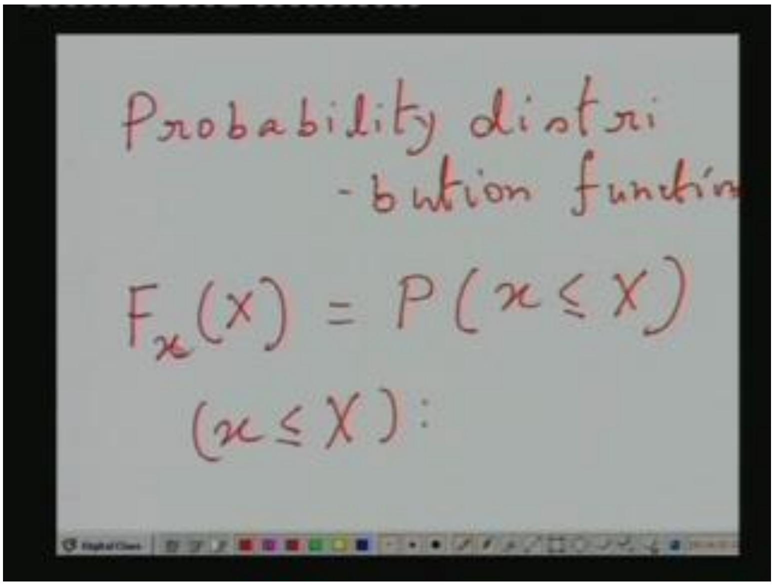
Probability & Random Variables
Prof. M. Chakraborty
Department of Electronics & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture -5
Probability Distributions and Density Functions

So, today we will be discussing probability distribution and probability density function very important concepts.

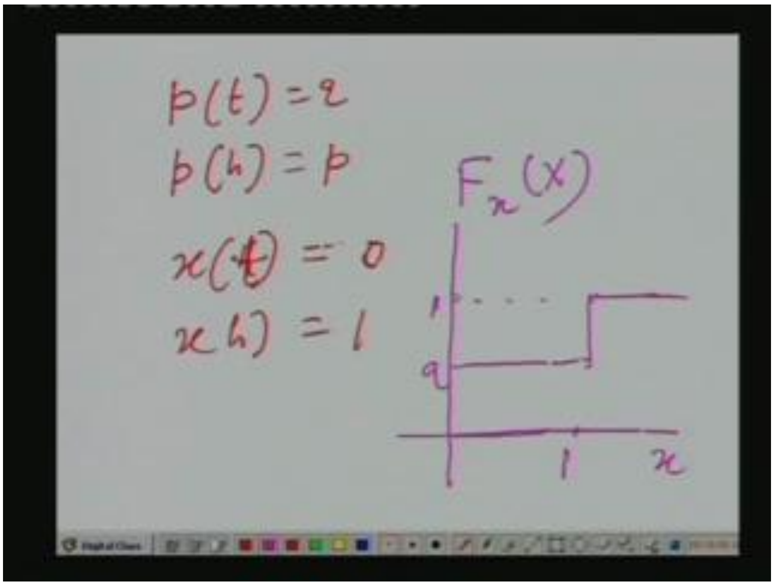
(Refer Slide Time: 01:04)



First we start with this is probability distribution function, you know it is written like this it means the probability of this. Let me explain X here is a real number in this case and this corresponds to an event what event I mean event, for which this value I mean this variable x takes values all either equal to X, X or less than X less than equal to X.

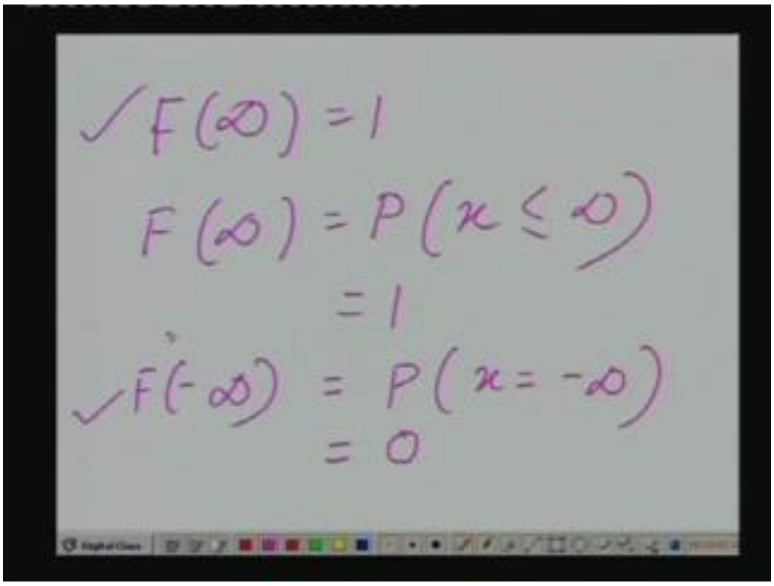
Now, probability of that is called the probability distribution function for the random variable x, if x X. So, the particular value that we have in mind is put here, this x denotes the corresponding variable. Sometimes I mean we can be little liberal here, we can omit this subscript and just directly write F X what you mean F x. Now, this will be clear if you give some example. What is here, the example of tossing a coin. Suppose, you are tossing a coin and the probability of tail occurring is a q that is p of tail t for tail is a q and probability of head is p further.

(Refer Slide Time: 02:56)



So, we have got just 2 experimental outcomes here t and h. Now, suppose there is a random variable x, which takes the value this way 0 and xt is 0 and xh is 1 xt is 0; just a minute, this is given now, what will be the corresponding probability distribution function. It will like this plot, this axis, this is for x and this x is this is for F. Now, as x goes from 0 to 1 you know from 0 to 1 Sx goes from 0 to 1. What is the corresponding probability in this range. Now, here x can be 0 or 1 I mean up to 1 if this is tail right. So, it is q and then it becomes 1 all right this is the probability distribution function. This probability distribution function has got some very important properties. Let us examine those properties.

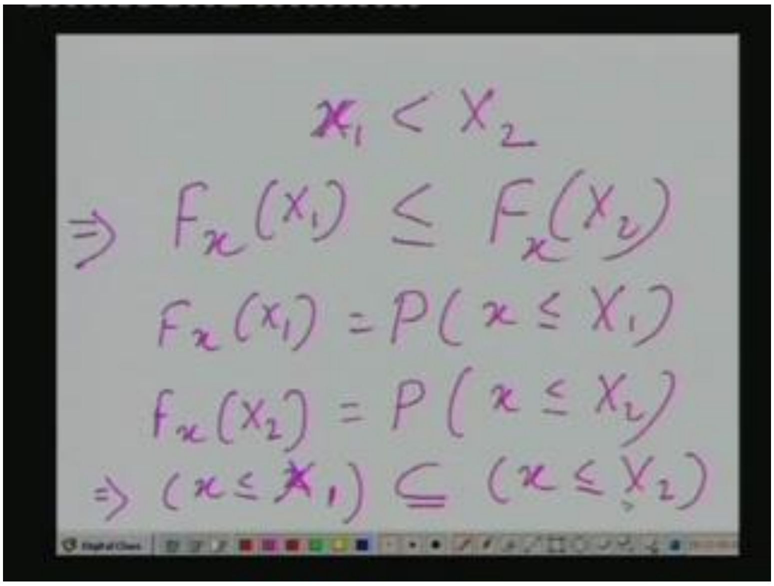
(Refer Slide Time: 06:10)



Firstly F infinity this is equal to 1. Now, why F infinity equal to 1, it is, because F infinity means P. Now, you see by the definition of a random variable for no experimental outcome externally can take the value infinity, because the corresponding probability was 0, which means this event includes all experimental outcomes, because all of them are taking finite values, for all of them existing finite values right.

So; that means, this come this means the total set or total event, for which the probability is one. So, this is equal to 1; on the other hand how about F minus infinity. Now, F minus infinity means; actually since we cannot go to the left of minus infinity this means px equal to minus infinity again this corresponds to no experimental outcome. So, this is an impossible event for which this is equal to 0. So, F infinity is 1, this is 1 property another property is this F minus infinity is 0. Next if you consider, things like this you know X1 less than equal to X2.

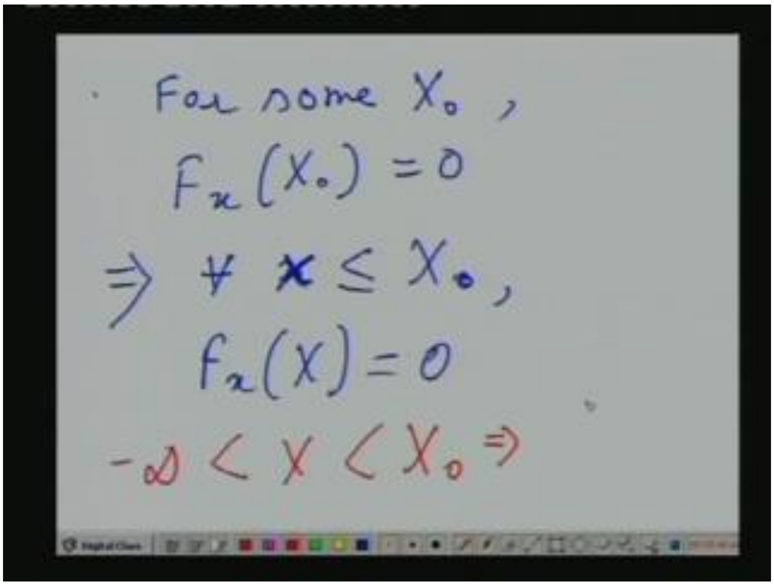
(Refer Slide Time: 08:11)



Then or whether this is less than equal to Fx 2; that means, the probability distribution function is actually a non decreasing function, it can remain steady or study or it can only go up. Now, this follows easily (( )) what is this Fx, x1 this is nothing but P and Fx. So, in this (( )) P, but you see you have given with this, which means; this set it is actually content are may be same in this set, because x1 is less than x2. So, set of outcomes for which, x is less than equal to X 1 that must be content, in this set that is set of all outcomes for which x is less than equal to x2.

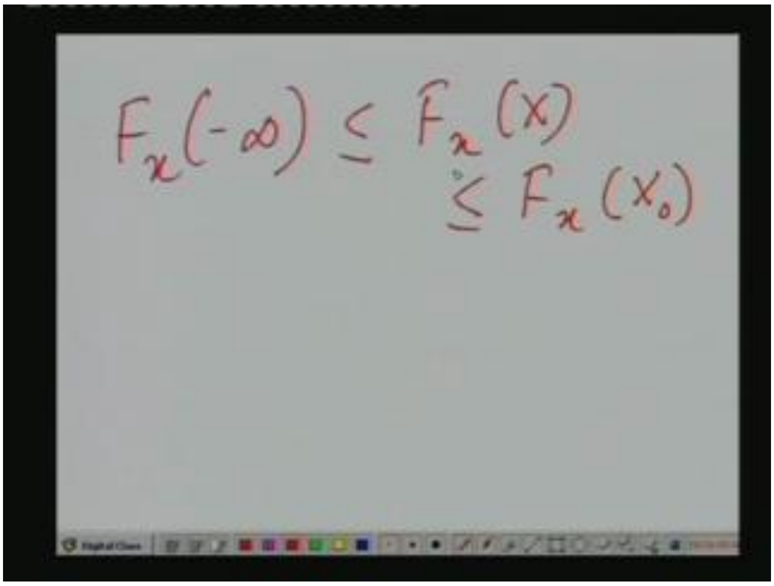
Now, we have already seen earlier, in the context of probability that for this event. If this event is content in another event then the corresponding probability is less than equal to if, it is perfectly content the corresponding probability of that event is strictly less than the probability of the other event, but since it is it can be same also. So, it means that this probability is less than equal to his probability which means; Fx x1 is less than equal to Fx X2, which means the probability distribution function is a non decreasing function it is very important all right.

(Refer Slide Time: 11:32)



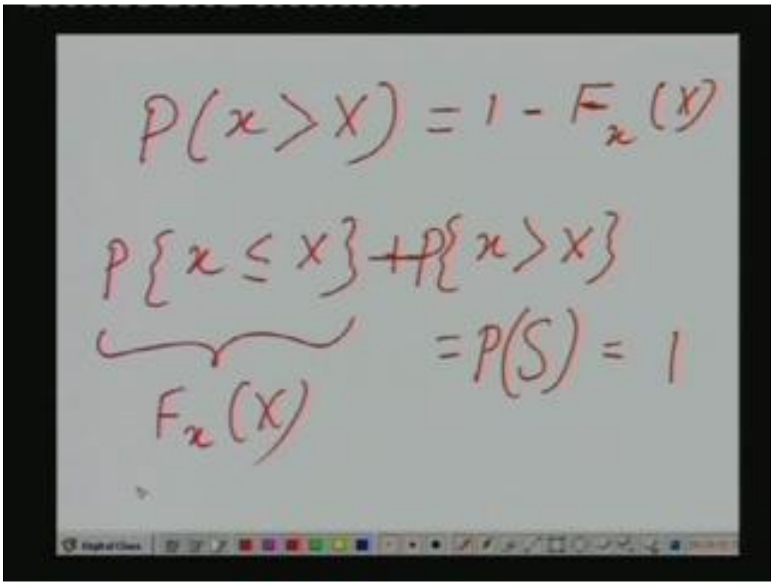
Next suppose, for some X0 consider, some number X 0 if it. So, happens that for some X0 Fx X0 is equal to 0, then it implies that; for all for all x less than equal to X0 for all x less than equal X0 also or let me, put this was 1 make it for all X less than equal to X 0 also we should have this. Again this comes from the previous theorem that: you know previous 2 theorems that. Firstly, F of minus infinity Fx minus infinity is 0 and the probability distribution function is a non decreasing function, because if you consider, any x which is say less than less than X naught. If it is equal to X naught this is already given. So, if it is less than X naught then you have got this thing right, but since probability distribution function is a non decreasing function this implies, but this is 0 and this is 0.

(Refer Slide Time: 13:26)



So, this has to be 0 all right. So, this is (( )) number 3.

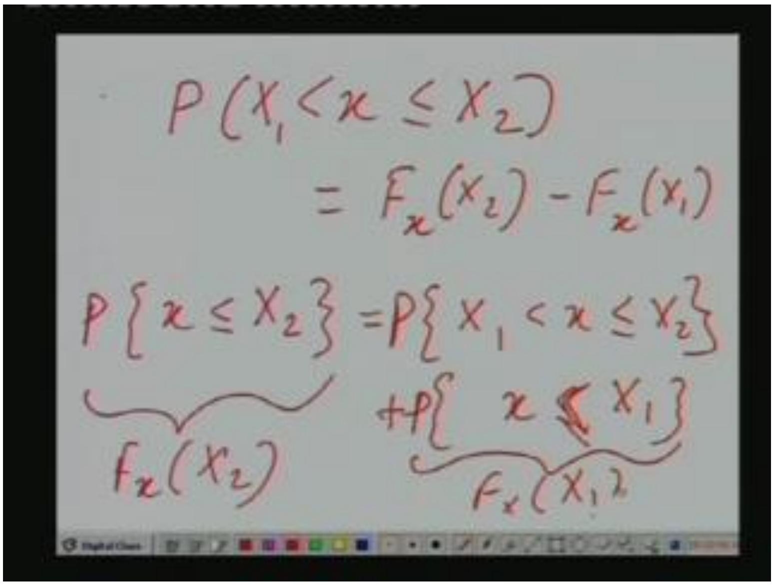
(Refer Side Time: 13:59)



Next you consider this. The probability of this random variable X taking value greater than some X, where X is any real number that is equal to actually 1 minus Fx X. Now, this is quite easy you know consider, these events, and this events, this event these are mutually exclusive and this means this means a total set actually. This means the total set say S and since, they are mutually exclusive and since, the probability of is total set is one; that means, P of this and P of this equal to this, because they are mutually exclusive.

So, P of this total set can be written like this, but since, is a total event that the most certain event it is probably is 1. So, this follows immediately, because this is nothing but Fx X. So, this is equal to 1 minus Fx Fx x. So, actually these 2 are mutually complimentary, probability distribution function; for a particular number X and the probability that the random variable takes values higher than X.

(Refer Slide Time: 16:53)



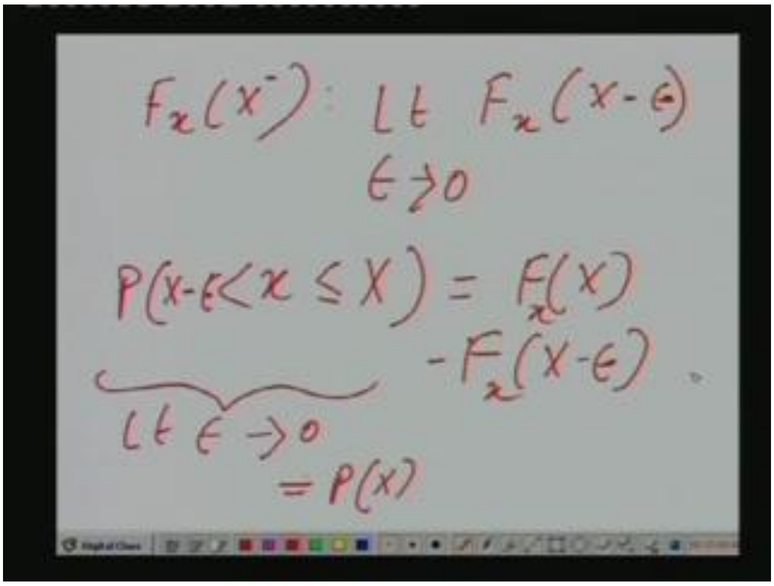
Next, probability of this giving 2 such numbers X1 and X2. What is the probability of this. This is equal to actually, again these (( )) very difficult consider, this set x this can be written as sorry is strictly less than. These are 2 disjoint sets in 1 case x lies between X1 and X2 another case x lies below X1 and their union gives raise to this set.

So, probability this will be what probability; this will be effecting the probability since, they these are disjoint this probability can be written as the probability of this plus probability of this all right. But, this is equal to may be; I make a little change here, at this stage actually, I am making certain change here x is greater than X1, but less than equal to X2 greater than X1, but yeah less than equal to X2 then this formula will be valid actually.

For that first I consider, this set x less than equal to X2 this you can decompose like this: x greater than X1 less than equal to X2 and x less than equal to X1. These are mutually disjoints. So, probability of this is equal to probability of this plus probability of this right, but this is equal to Fx X2 and this 1 is equal to Fx X1.

So, things getting bit dirty here, so let me erase and. From this follows; obviously, the this probability which is this equal to Fx X2 minus Fx X1; these are very important actually. These all these might appear to be trivial, but actually these are the you know very useful as you go on go along you know, because actually to find out some probability of some event occurring in some range in all that you know all this properties come very handy.

(Refer Slide Time: 21:03)



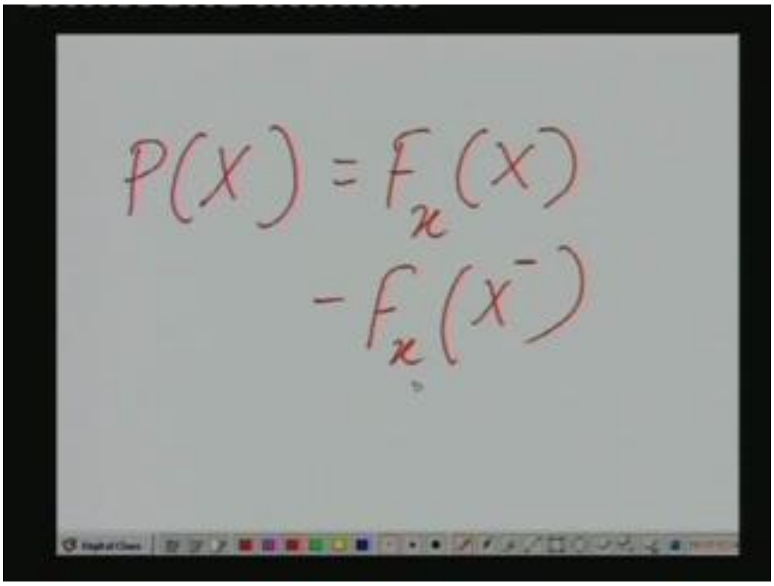
Then another thing, I defined this; this means, that if you take this probability distribution function not at X, but slightly to the left by some you know its slightly less by a quantity, which is infinite through infinitely, but positive real called epsilon and then if you take this limit epsilon tends to 0. So, that; that means, you are basically approaching this probability distribution function from left then the limit that you get I call it Fx X minus.

Now, see 1 interesting thing that; using the just previous theorem. We can always right this thing that P of this is equal to this comes from the previous theorem, because this number is greater than this sorry this should be F actually and this is epsilon. Now, see the interesting thing, if on the left hand side, I take a limit of epsilon tending to 0 then what happens that is I take a limit on this I take this limit.

Obviously, as epsilon become smaller and smaller this lower limit becomes closer and closer to X. So, the gap between X and this becomes smaller and smaller. So, in the limit x coincide with X. So, this when I take the limit with epsilon tends to 0 these goes to px and we have seen that if on the probability distribution function.

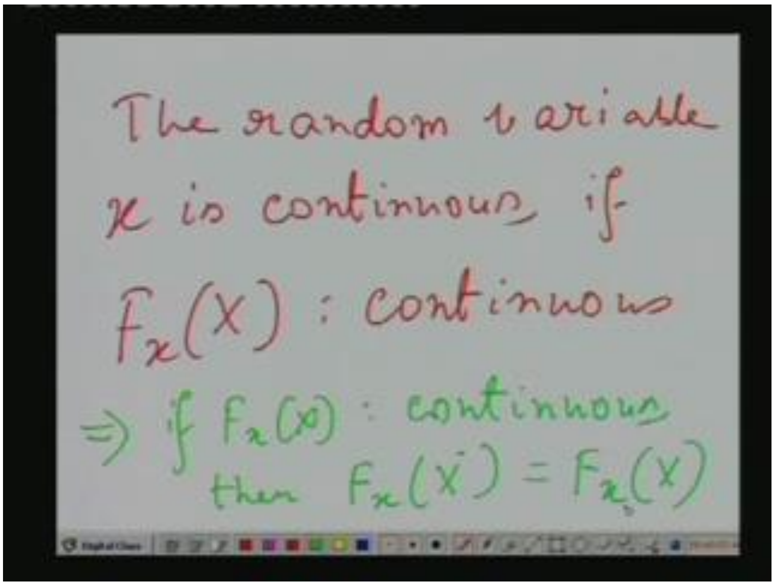
We take the limit from the left hand side that is (( )) epsilon tends to 0, we have a notation for it Fx X minus; that means, P of X P of X there is X taking exactly the value X is nothing but probability distribution function up to X minus the left limit of the probability distribution function for the point X.

(Refer Slide Time: 25:02)



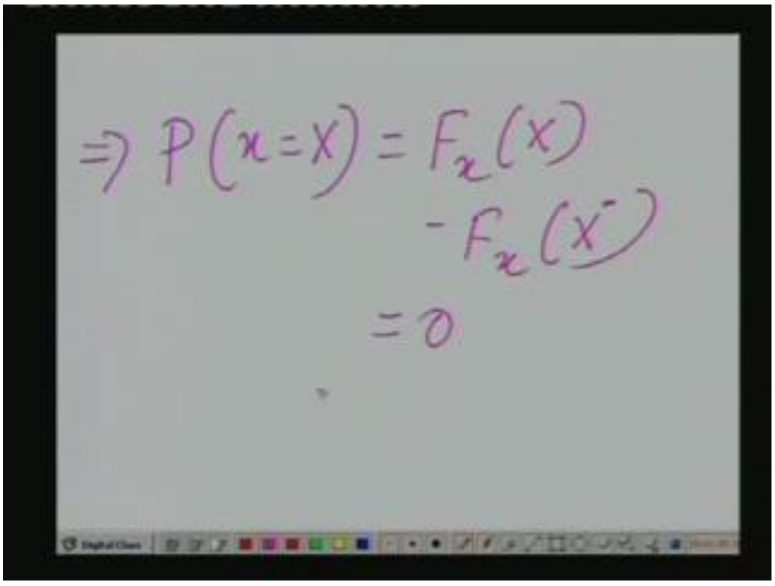
That means px all right. Now, this probability distribution function can be continuous or can be discrete type. If it is continuous that is in that case we will call it, we will say that the random variable itself is continuous that is the…

(Refer Slide Time: 26:11)



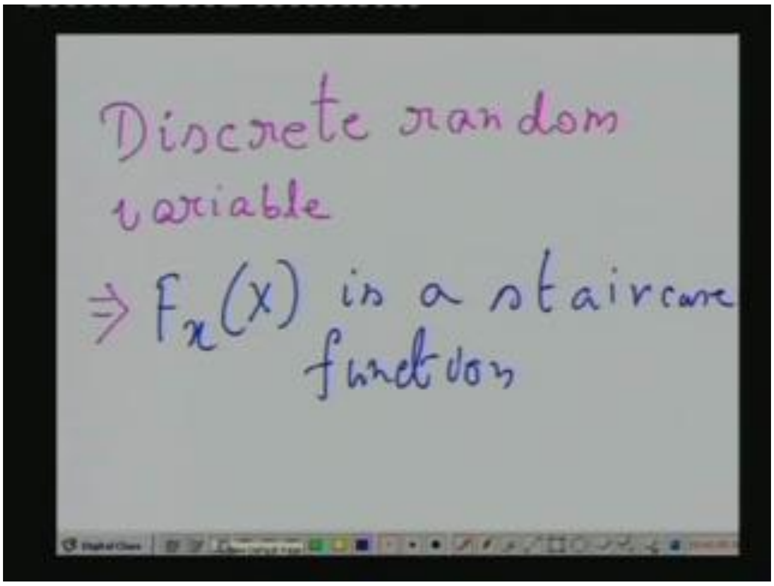
If corresponding, probability distribution functions continuous. If it is continuous let it means; this means, you see if this is continuous then this limit will be same as this comes from the definition of limit and continuity. The limit from left should be same as the value itself what does that mean then it means; if the random variable is continuous. And therefore, this probability distribution function is continuous up to here. We have got this the left limit of the probability distribution function at X is same as the value of the probability distribution function at X.

(Refer Slide Time: 28:12)



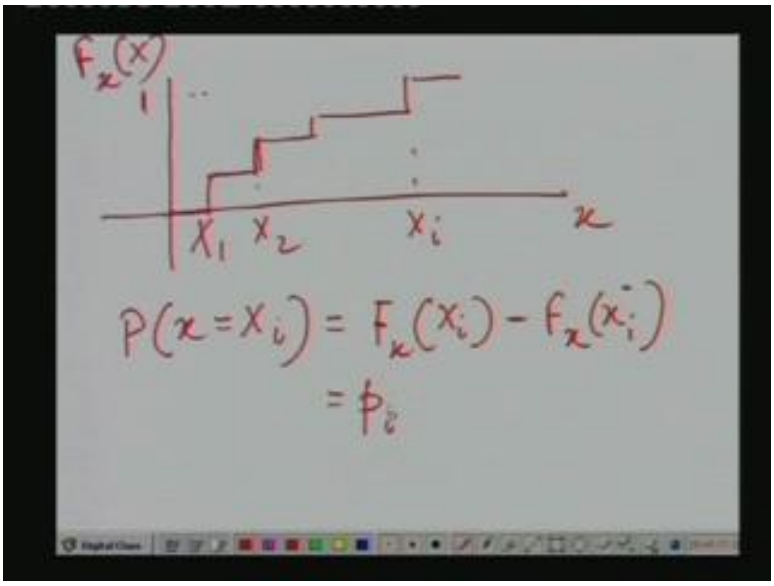
This would imply from our previous theorem that, probability of x taking the value X exactly is the probability x taking X that by our previous theory has been this, but seems these 2 are same you know this becomes 0. So, this is the property of continuous random variables, but there is another random variable called discrete random variable.

(Refer Slide Time: 28:48)



Here, the probability distribution function is a staircase type of function. I will give an example.

(Refer Slide Time: 29:36)

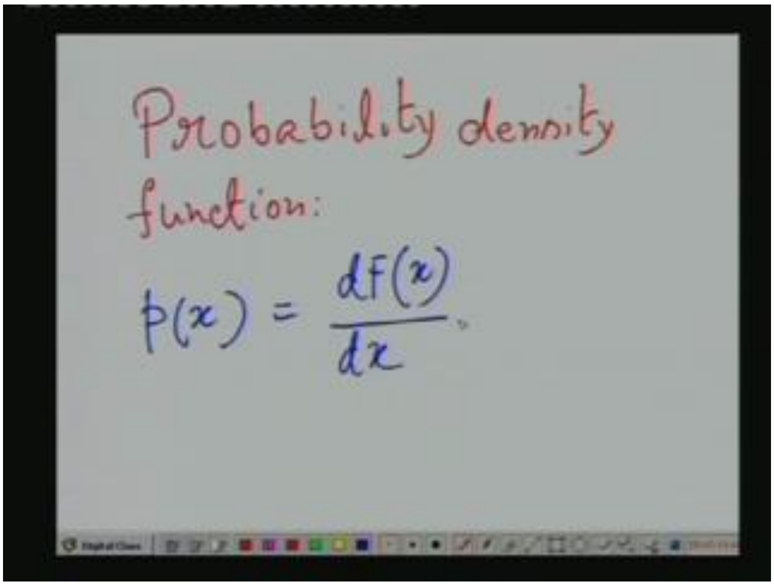


These are the points of the discontinuities say: X1, X2 say this is: Xi right, this axis is your probability distribution function, this will be the case of a discrete random variable. In this case you see the function itself is not continuous there are discontinuities**.** What happens then at these discontinuities. You see up till this point that is: just about just at the point, when we are about to go to X2 till that time this probability distribution was giving simply the probability of 1 event and then at X2 there is a jump there is a discontinuity right.

So, what is or instead of X2 say consider, Xi if I say what is the probability of x taking Xi exactly this biotheory would be this. Now, you see what is the meaning of this, Fx Xi means; what this is a distribution function; that means, probability of take X is taking value up to Xi that is: either Xi or any value up till this; that total probability is this and what is this left limit. Left limit is nothing but the probability of you knows, this variable X taking value up till Xi, but excluding Xi.

So, this difference will be exactly equal to the probability of X taking Xi you can even call it Pi. These are the case of a discrete random variable. Now, from probability; probability distribution function, we actually go to something more interesting and more widely used that is called probability density function.

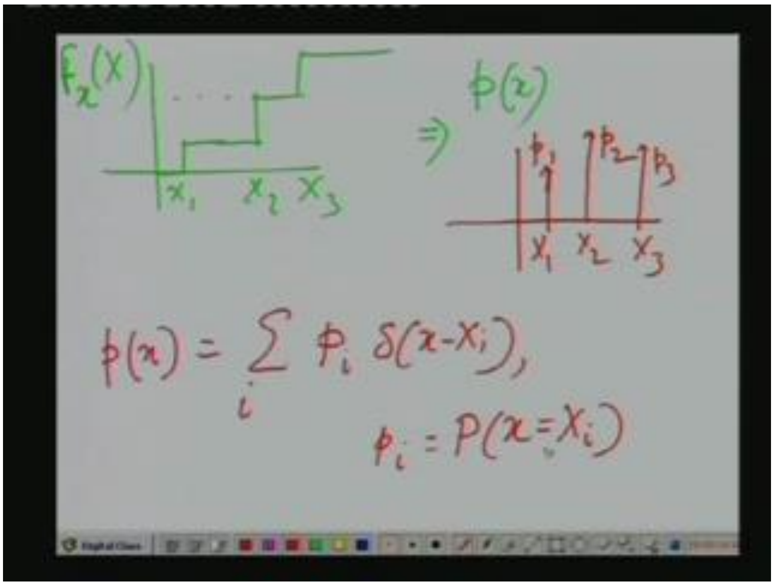
(Refer Slide Time: 33:18)



Here, sometimes we denote by f sometimes by p you know my notation should be whether P, because P is stands for probability actually. So, I preferred P, Px it is nothing but the derivative. Here, for simplicity I am just writing it as a function of x this probability distribution and deriving it with respect to x this is called the probability density function; 1particular, 1 interesting thing you can see that suppose, this is the continuous function then of course, we have got px as continuous.

But suppose fx is a staircase kind of function that is: the random variable is discrete random variable, in that case this derivative will be 0 everywhere except at the discontinuities and the at the discontinuities there will be impulses right (( )) functions impulses is multiplied by the probability of x taking you know values at those discontinuities that is px.

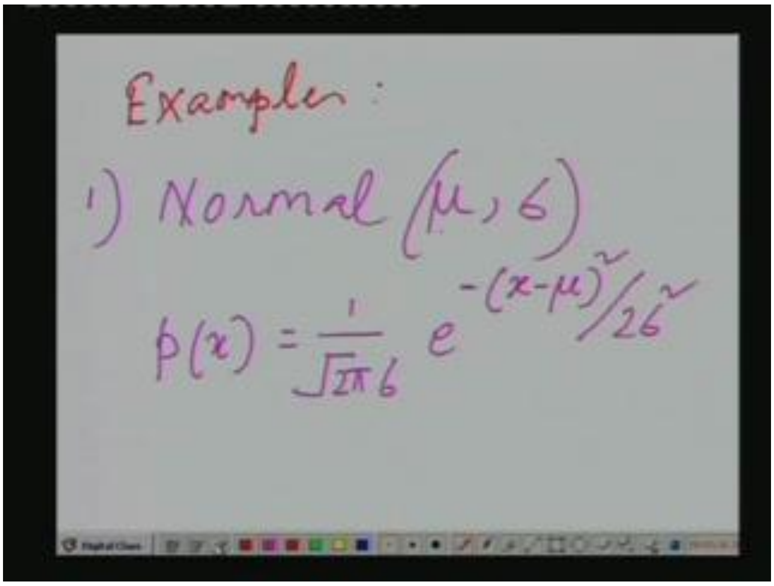
(Refer Slide Time: 35:38)



For, a function like this. Let us again, once again draw function like this: you know X1, X2, X3 and things like this; this for the probability distribution function. Corresponding px should be like this, you know have 1 impulse here a strength will be p1, where p1 is the probability exact probability of X taking value equal to X1 then there will be impulse here then another impulse and like that p3. This is again X1, this is again X2, this is again X3 and like that.

So; that means, you can write in general px is nothing but summation of pi delta x minus Xi and this is what sum i. And what is pi, where pi is probability of x taking value equal to Xi. This is for the discrete case. So, remember for a discrete random variable the probability density function actually is a linear combination of some impulses. It will be interesting to consider, some standard probability density functions like you know normal; normal distribution or you know I means uniform and things like that.

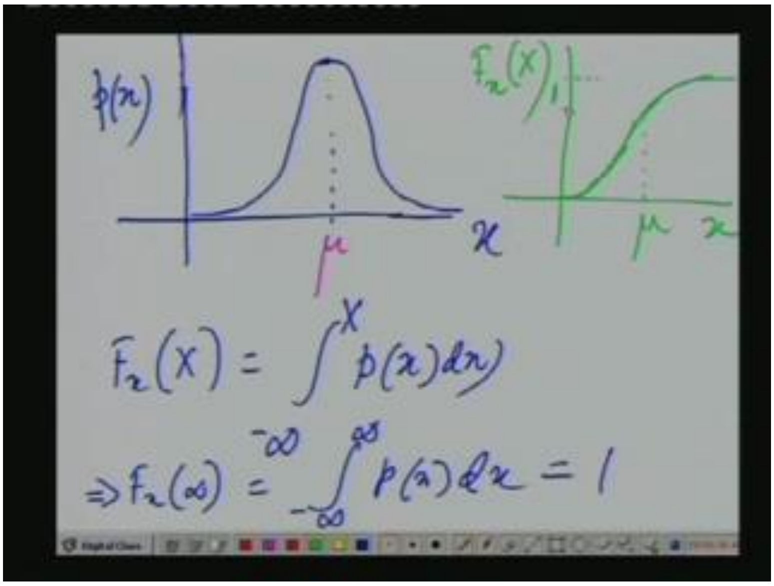
(Refer Slide Time: 38:19)



Number 1: normal we call it mu these also Gaussian sometimes, it is called normal sometimes it is Gaussian is formalize px. It might look complicated, but is 1 of the most widely used density function especially in communications and thing like that. These 2 parameters mu and sigma they have their respective interpretations expect the meanings which will study later; mu is called the mean and sigma is called as the standard variation, or sigma square is call the variants of the random variable x.

Now, since we have been define these things I am not getting into that, but this is a typical form you can see 1 thing that since, there is a squared term here this will be symmetric I do not mean that is: where that x minute mu is positive or negative, we will get the same value right and when x equal to mu then only, it will have its (( )) value then as this difference increases since it is coming with a negative power this decreases. So, it will be having a peak value somewhere, and at x equal to mu and then as you go to the right and to the left of mu it false down.

(Refer Slide Time: 40:59)

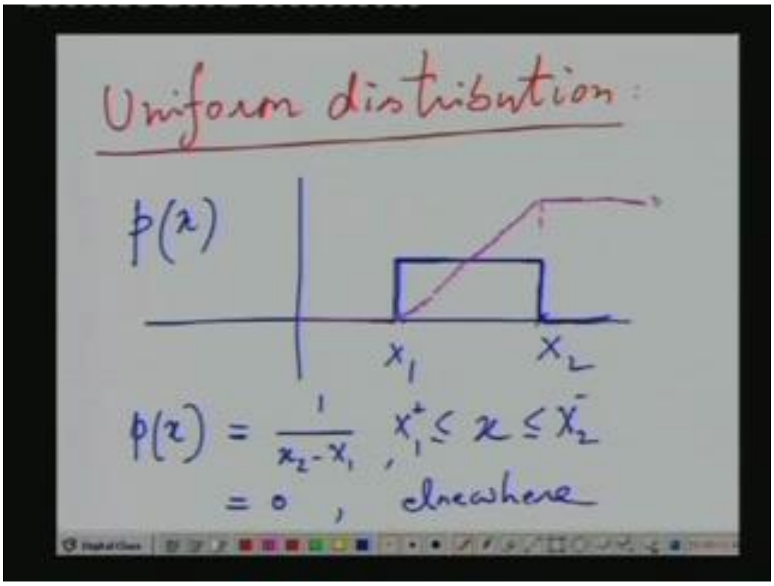


So, typically it will be like this. Is actually a bell shaped curve is occurring at mu. This is px, this is x and 1 thing you have seen that what is px after all it is the derivative of fx which means this nothing but integral of px, dx from minus infinity to X; that means, Fx infinity is nothing but integral of this from minus infinity to infinity and since, that is equal to 1 you know we have this thing this equal to 1 which means; this probability density function has a total area under it equal to 1.

So, when you integrated from minus infinity to infinity, the total area should be equal to 1 that is a case here. For this case actually, the corresponding distribution function would be like this, to get the distribution function you have to go on integrating right and as you go on integrating area increases. So, distribution function will go up and finally, when you have come to this setup. If the values close to this after that; if it is some increase at x the value of this probability this is the function has fallen down, so much that the probability distribution function does not increase much it (( )) of remain study.

So, initially there is a growth like this, your mu will be somewhere at here, but as x increases finally, this approaches 1. This is for the Gaussian case or normal case. There are many other interesting example, today I will consider one more uniform and few others like: binomial Poisson (( )) and all that we will consider, in the next class. So, another example again quite widely used I would say of a probability density function is that of uniform distribution.

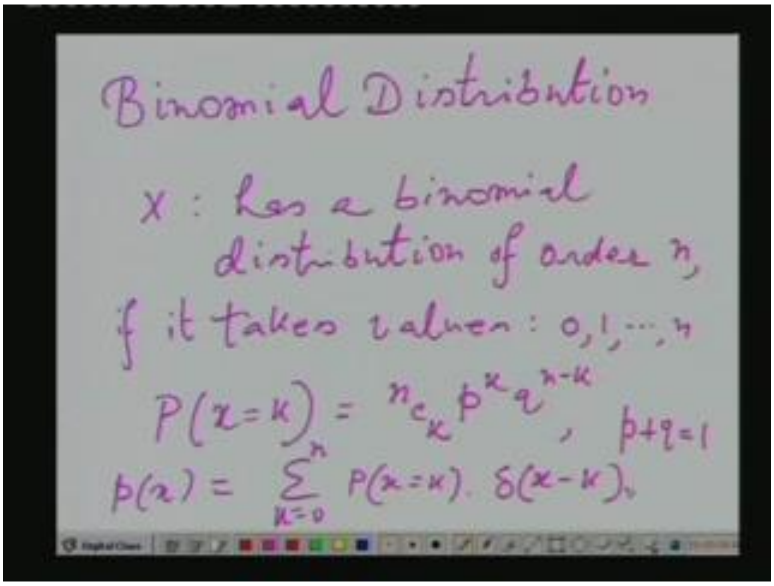
(Refer Slide Time: 44:23)



This means; in this case sorry this is something like this, you know take 2 numbers X2 X1 within this density function has a constant value, but to the right of X2 and to the left of X1 its value is 0. Obviously since, the total area under this curve has to be equal to 1, then we can find out how much would be height, it is width is X2 minus X1. So, height should be 1 by X2 minus X1.

So, here I am whether writing minus and plus, because there is a discontinuity here and this is equal to 0 elsewhere. So, this is for the probability density function and corresponding probability distribution function will be what we have to gone integrating it. What will happen if you integrate it. In up till now, up till point 0 that as you go on integrating the area increases finally, at this point area it has maximum and that is equal to 1 and there it remains steady right. So, for the probability distribution function should be like this, is going up linearly like this. So, this is for the case of uniform distribution. Consider a discrete case, suppose there is a random variable (( )) course here, what is called binomial distribution.

(Refer Slide Time: 47:40)



Suppose, there is a random variable X we call it that it has a we say that it has a binomial distribution; has a binomial, if it takes the values only this either 0 or 1 dot dot dot dot up to n with the corresponding probability like this. These are the probabilities this is nck p to the power k q to the power n minus k; 1 thing you see where of course, p plus q this is equal to; see if you really sum up that is p of x equal to 0, then x equal to 1 up to x equal to n you know and add them then you get a binomial series p plus q whole to the power n. And since, p plus q equal to 1 you get the value 1 which should be satisfies, because x can (( )) this or this or this x cannot take anything else.

So, total probability should be equal to 1. So, this is called binomial distribution of what are n. Of course these are only the probability values, where x I mean probabilities for x taking either value 0 or 1 or 2 and things like that, but when it comes to the density then as I told you earlier you need to add like this, so this for binomial distribution. Now, there are other distribution also, you know Poisson distribution, then Rayleigh distribution and which are of course, very useful in communication or you know time phase modeling and weather prediction and things like that.

Since, we do not have much time, now we will concept them in the next class. So, in today’s class just to sum up what we have done; we have started with probability distribution function, its definition cauterization we discuss we gave some example. For this probability distribution function, we discuss various prosperities like you know I mean the probability distribution function at X equal to infinity has value 1 at X equal to minus infinity has value 0, then probability distribution function is a non decreasing function that is, it can only go on increasing or it can remain steady it won not go down and then some more properties were discussed.

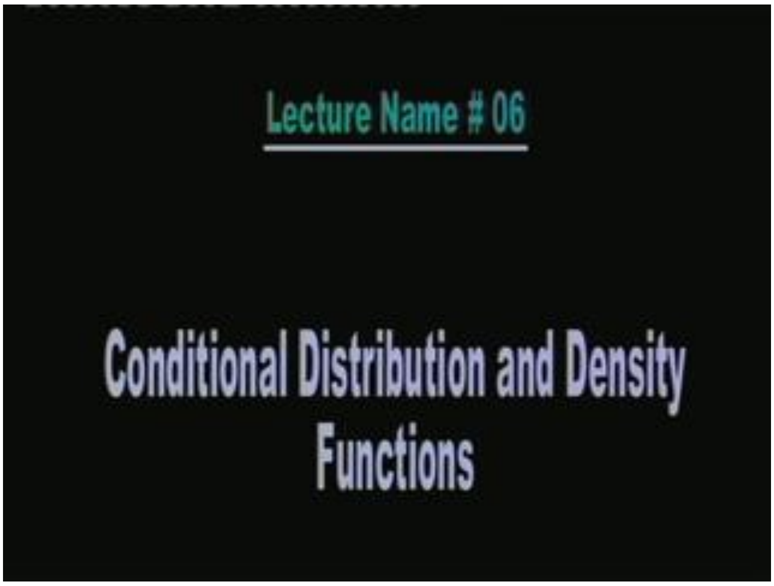
That is if there is a random variable which takes values between 2 ranges you know 2 figures X1 and X2 these are the probability that the random variable will remain between X1 and X2 can be described as a difference between the 2 difference between the probability distribution function at X2 and X1. We discussed what it meant by continuous random variable, what it meant by discrete random variable.

Then we discussed, we gave a definition of the probability distribution function and gave certain examples like: Gaussian and Gaussian also is called normal distribution, then gave the example of uniform distribution and conclude with binomial distribution. So, the next class, we will be taking few more examples of this distributions we will carry out some examples also and then we will go into conditional probability density and distribution. So, that is all for today's lecture.

Thank you.

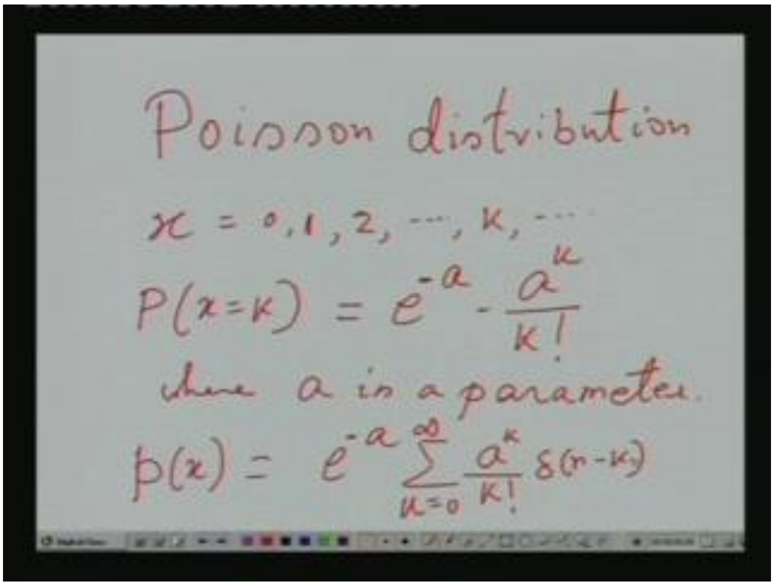
(Refer Slide Time: 52:50)

(Refer Slide Time: 53:00)



In the last class, we discussed the probability distribution, and probability density function and discussed the properties of probability distribution function. And then we consider, some examples of probability density function like you know I mean normal that is Gaussian density, then binomial distribution that was the distribution, and we consider the corresponding density also probability density function also then we considered binomial (( )) we considered, uniform probability density and all that. We in today's class, we start from that we consider one more probability distribution function, which is very useful especially in queuing theory and network related problems that is called Poisson distribution.

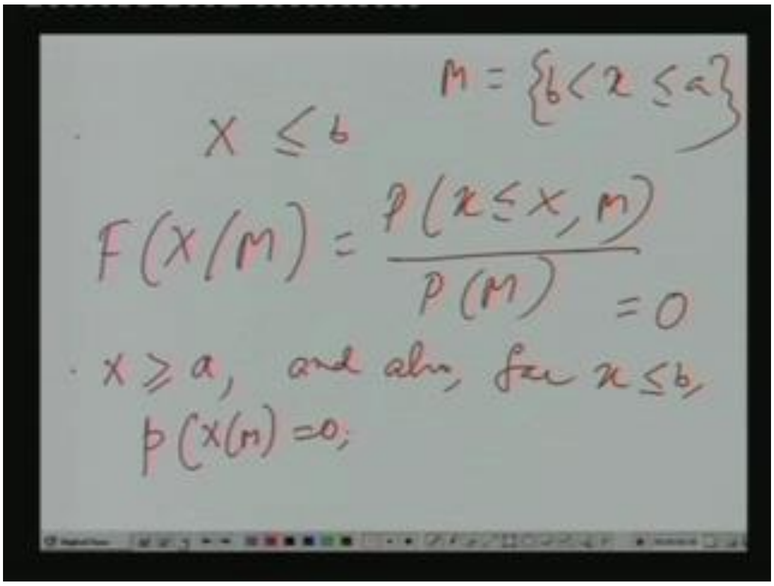
(Refer Slide Time: 54:06)



Here x, it takes values like you known 0, 1, 2 dot dot dot dot K dot dot dot dot. The probability of x taking k, will be given as e to the power minus a into this is into a to the power k by factorial k, where a is a parameter; question is can a be negative. You see a cannot be negative is, because if you consider, this in a negative a odd powers of k what values of k that is odd powers of a here, say a to the power 1 or a to the power 3 or a to the power 5 and all that. And obviously, if a is negative the entire thing becomes negative right at probability cannot be negative.

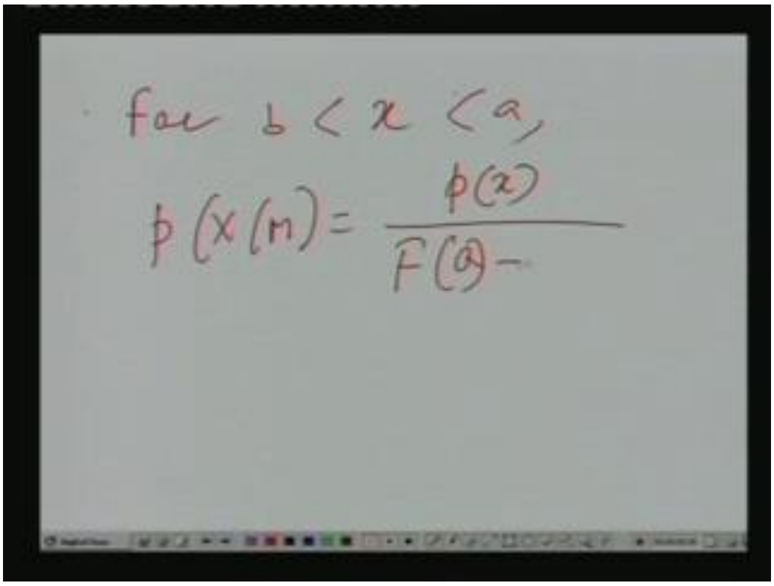
So, from the in the definition itself, it is implied that a is a positive number. And depending on the a that; we choose you get 1 particular distribution or other right. So, if that we... So, these are corresponding probability density function. So, this just a probability is not (( )) probability distribution, it is just the probability of x taking k. So, corresponding density function as you discuss the other day will be what, they will be impulses at 0, at 1, at 2, at k and like that and each impulse a that impulse at k and in the other case other range.

(Refer Slide Time: 56:22)



In X less than equal to b. In that case what is happening our m is given as these and what is the intersection here, X is less than equal to b whereas, m corresponds to those x which lie in this range greater than b less than equal to a. So, intersection is 5 mccabe thiele method, for which probability is 0. So, clearly this will be 0. So, corresponding density will be what; when in this probability distribution function is 1 that will be 0, when this is 0 that will be 0 and in between in between, it will be just derivate of this and then derivative will be what; that is: for X greater than equal to a and also for x less than equal to b.

(Refer Slide Time: 57:52)



P of x by m. And for x greater than b less than a p of x by m, will be p of x, because we are deriving it was Fx minus Fb. If you derive you get px only and in the denominator you get Fa minus Fb. So, that is all for today. And we have covered enough of this topic on random variable and probability density and distribution functions and their properties.

Then condition probability density conditional distribution. We have just to consider, what is called total probability distribution, total probability distributional all that. We will take of some few examples, in the next class. And then we will proceed to what is called function of a random variable. So, then we have been dealing with only a random variable, then we have to consider function of a random where you likely. Now, I mean if x is random variable then and its probability density is given, but then x square also is a random variable, what will be its probability density function and likewise. So, that will be done in the next class.

Thank you very much.