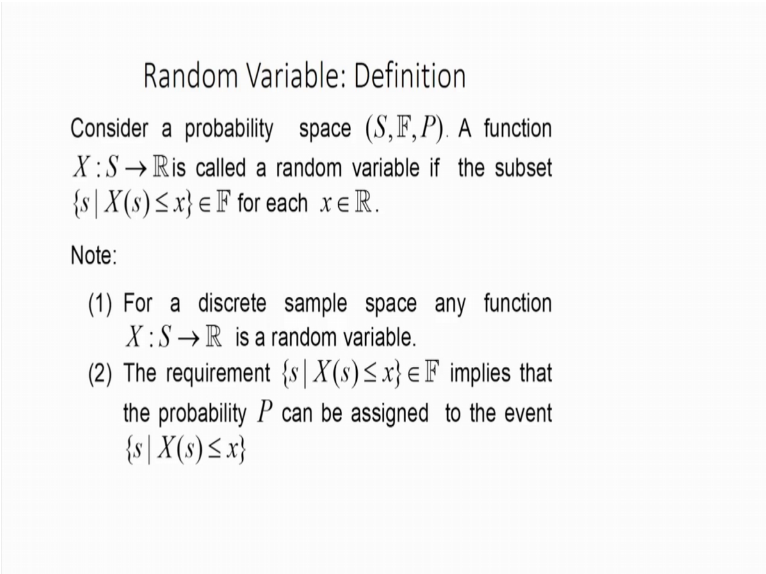
**Advanced Topics in Probability and Random Process**
**Prof. P. K. Bora**
**Indian Institute of Technology, Guwahati**

**Lecture - 02**
**Random Variable**

In the last lecture we discussed about Basic Probability Concepts. In this lecture we will introduce the concepts of random variables and how to characterize them a random? A variable is basically a number assign to the sample point so that we can numerically analyse the data. Let us see the definition of random variable.

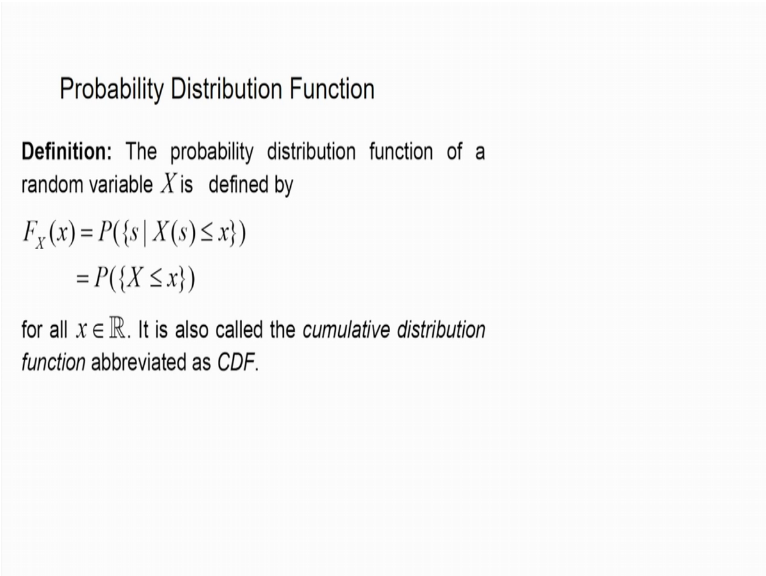
(Refer Slide Time: 00:53)




Consider a probability space S, F, P. A function X from S to R is called a random variable if the subset s such that X s is less than is equal to small x belong to F for each x belonging to R. So, this is the definition.

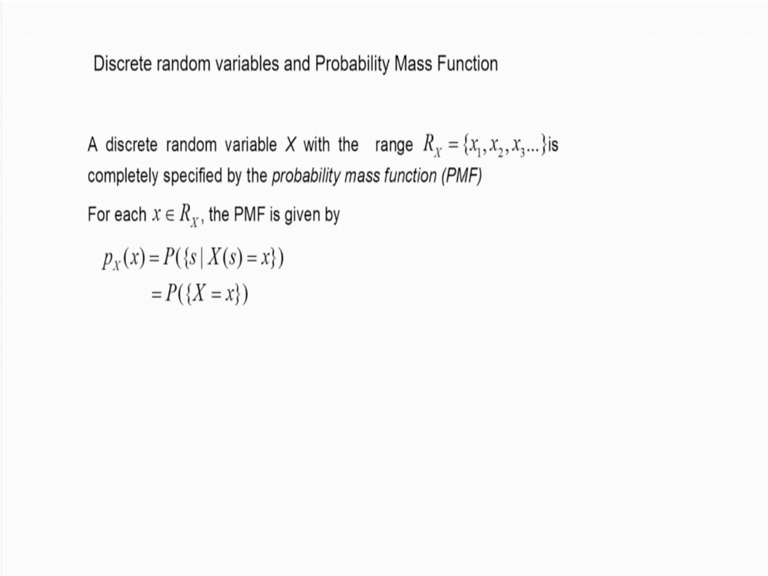
First of all the random variable is a function from the sample space S to the real line , but there is a requirement the requirement is that if I consider any point x on the real line we can define an event like this s such that X s is less than equal to small x. So, this event should belong to F, this requirement is only for uncountable sample space this implies that the probability P can be assigned to the event this event as such that X s less than equal to small x. Because we have define we have to define a probability this probability should be well defined. Otherwise for a discrete sample space any function from X to R can be considered as a random variable.

(Refer Slide Time: 02:07)



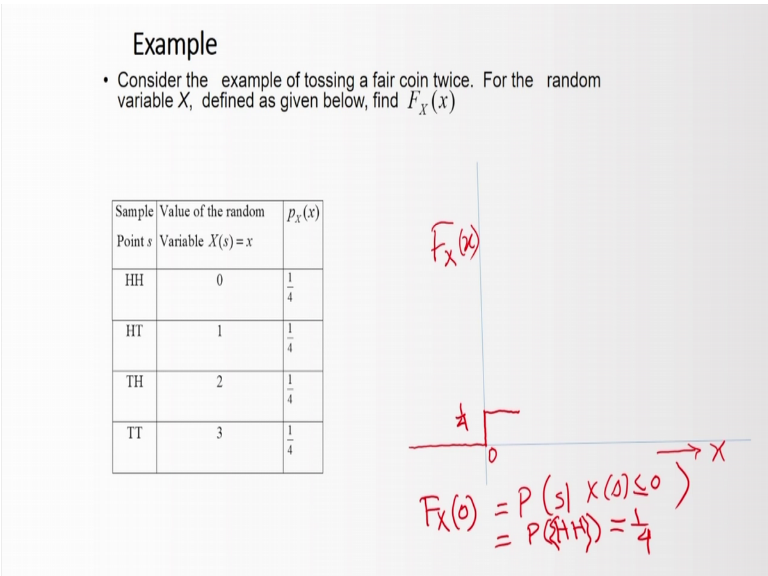

So, we have defined the random variable now how to characterize a random variable, First of all we will define probability distribution function. Now we know we have defined an event s such that X s less than equal to small x this is the event. So, the probability of this event is known as the probability distribution function at point x F x of x is equal to probability of the event consisting comprising of s such that X s is less than equal to small x. And for notational simplicity we simply write probability of X less than equal to small x this is also called the cumulative distribution function therefore, it is abbreviated as CDF.

(Refer Slide Time: 02:58)




We will introduce also the concept of probability mass function consider discrete random variable with the range these are the values random variable can take. Suppose we enumerate them x 1, x 2, x 3 etcetera. Then the probability mass function that is denoted by p x of x, so this is capital x is for the random variable at a value x it is the probability that X s is equal to x. So, we have to find out all those x for which X s is equal to x, this also for notational simplicity we have written probability of X is equal to small x.

(Refer Slide Time: 03:42)



We will consider one example consider the example of a tossing fair coin twice. For the random variable x defined as given below find F x of x we have defined the random variable suppose tossing a coin the sample space can HH, HT, TH, TT these are the 4 members and we assign 0 to HH, 1 to HT, 2 to TH, and TT is assigned 3.

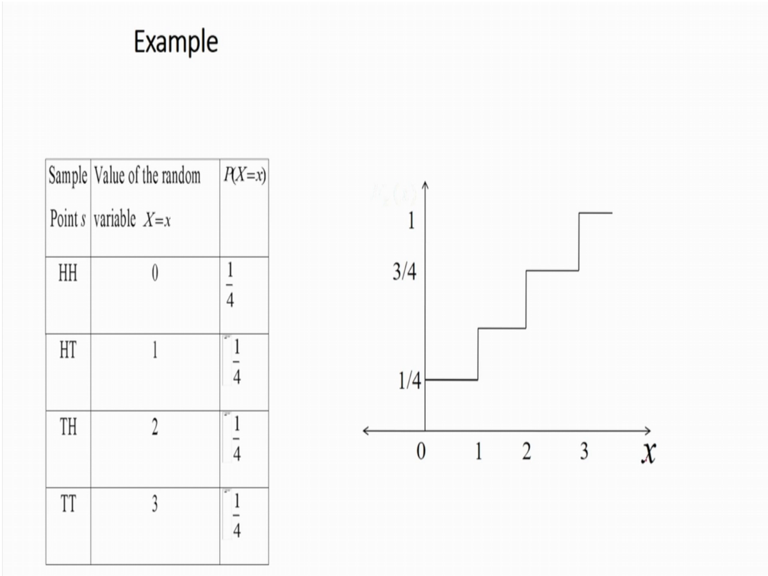
So, therefore, the value of the random variables are 0, 1, 2, and 3 these are these are discrete values and corresponding probabilities are given by suppose equally likely. So, probability of X s is equal to 0 that is HH, probability of HH that is 1 by 4, similarly probability of x is equal to 1 is 1 by 4, x is equal to 2 is 1 by 4, x is equal to 3 is 1 by 4.

So, we have to draw the PDF, so for that this axis is our x and this is my F x of x. So, let us see what happened because we have to define for all values of x. So, when x is less than 0, suppose in that case there is no event for which axis less than equal to 0 less than 0 so that way up to 0 it will be 0 only.

If I consider suppose F x of 0 suppose I have to find out F x of 0, so F x of 0 that is probability that s such that X s is less than equal to 0. So, what will be this event? We know that this is for which X s is less than equal to 0, so less than equal to 0 means there is only one s that is HH that is probability of HH, so that is equal to 1 by 4, so this way point we will have 1 by 4.

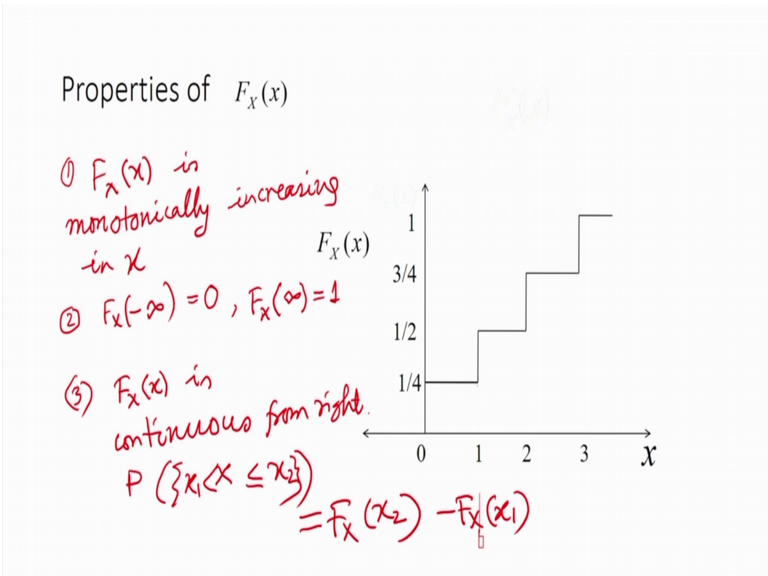
Similarly if I consider up to 1, so for all the probability will be same. So, we will have 1 by 4 like this, so that way we can complete this CDF plot. Let us see how these CDF plot look like.

(Refer Slide Time: 06:20)




So, this is the CDF plot. So, x 0, 1, 2, 3 these are the values correspondingly those CDF are up to here it is 0, then there is a jump and then this will remain at 1 by 4 again at 1 there is a jump this will become half like that. And finally, this is 1 F at point 3 and then hence further it will remain 1. So, this is a CDF plot for this random variable discrete random variable given by these values. So, we have seen this plot and on the basis of this plot we can describe the properties of this CDF.

(Refer Slide Time: 07:09)



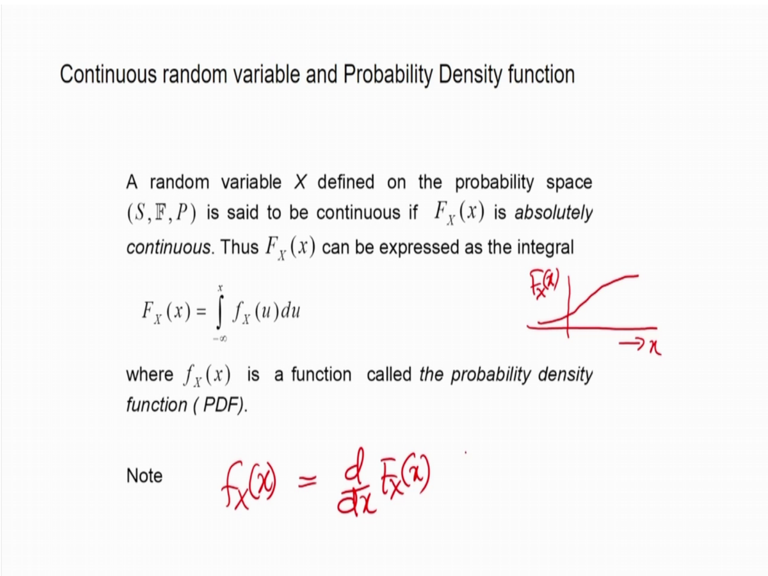
So, if I consider the properties of this CDF F x of x we see that generally this will be a non-decreasing function. So, number 1 is that a we can say that F x of x F x of x is monotonically increasing in x. So, x increases F x of x monotonically increases means it may remain constant or it will increase.

Number 2 we also observe that F x of minus infinity is if I extend here up to minus infinity F minus infinity this will be equal to always 0, similarly F x of plus infinity is always 1, number 3 we also observe that suppose at this point at point 1 what is the value of F x of x? That is the value because while defining this probability we consider capital X less than equal to small x therefore, at this point at point one the value is equal to half.

So, what does it mean? That F x of x is a right continuous function is continuous from right so; that means, now what is the value of the function at this point whatever we get by approaching from the right hand side that is the value. So, that way these are the properties and using the distribution function we can completely characterize the random variable we can find out any probability.

For example, probability that x lies between suppose if I say that x lies between some x 1 and x 2 ok. So, that probability will be equal to pro F x of x 2 minus F x of x 1. So, that way we can find out any probabilities involving this random variable x. We will define Continuous random variable and Probability Density function.

(Refer Slide Time: 10:06)

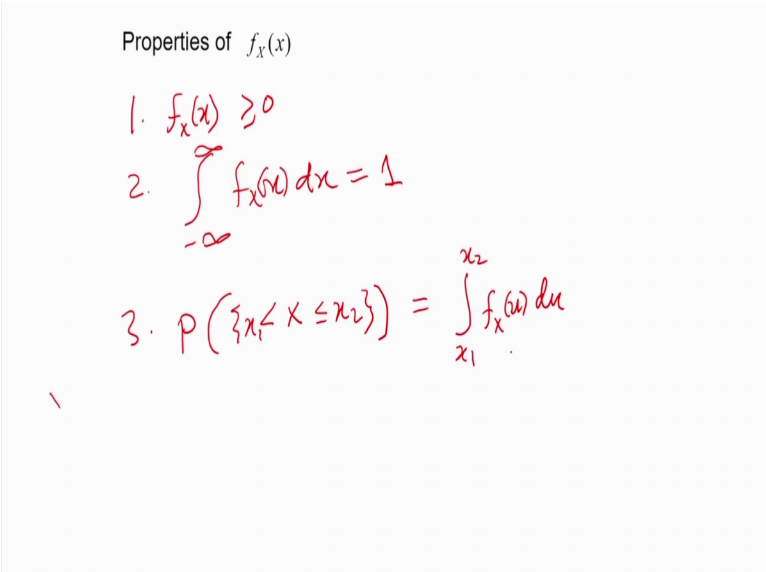


Now, what is a continuous random variable? If corresponding CDF is absolutely continuous. So, a random variable x define on the probability space S, F, P is said to be continuous if F x of x is absolutely continuous which implies that F x of x can be written in terms of an integral this integral from minus infinity to x small f x of u du.

So, this quantity suppose I have a CDF like this suppose this x F x of x. So, this CDF is like this suppose this is a continuous absolutely continuous; that means, this variations are bounded we will not go deep into this concept, but what is important is that F x of x can be expressed at the integral of this density function.

So, now we can write or we can tell that this f x of x is the probability density function it is abbreviated as PDF. Now for a simpler case was this CDF is differentiable function. So, in that case we can write F x of x is equal to small f x of x is equal to d d x of capital F x of x. So, PDF can be expressed as the differentiation or differentiation of the CDF. So of course, in that case we required that F x of x should be a differentiable function. So, we can derive some property of this PDF.

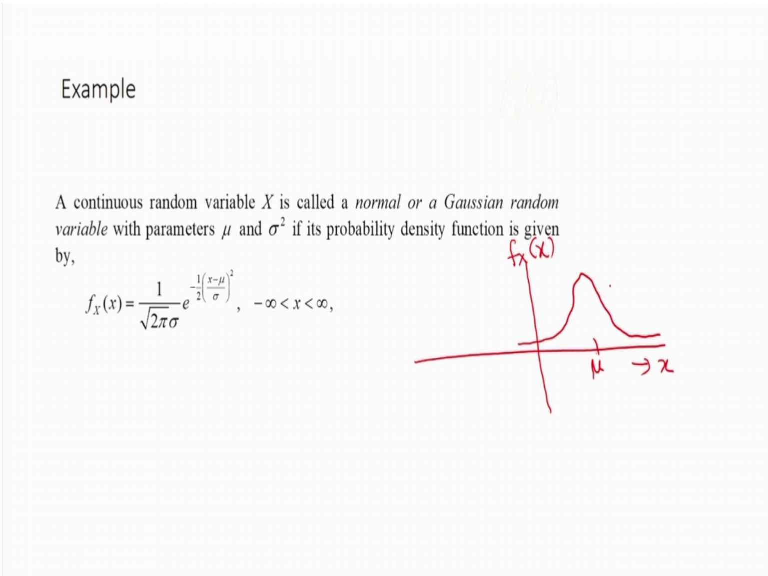
(Refer Slide Time: 12:17)




For example we know that CDF is a non decreasing function. So, if the PDF is differentiation of the CDF then we can easily show that f x of x is always greater than equal to 0. Number 1 it is a negative function, number 2 now the value of the CDF at the infinity is always equal to 1. So, from the definition we can easily say that this integration of F x of x dx from minus infinity to infinity that will be is equal to 1.

Also we can find out any probability of suppose the event any event suppose that event is x lies between some x 1 and x 2. So, this probability we can find out using the PDF from x 1 to x 2 integration the PDF f x of u d u. So, that way we can find out any probability involving the random variable continuous random variable x with the help of the PDF.

(Refer Slide Time: 13:42)

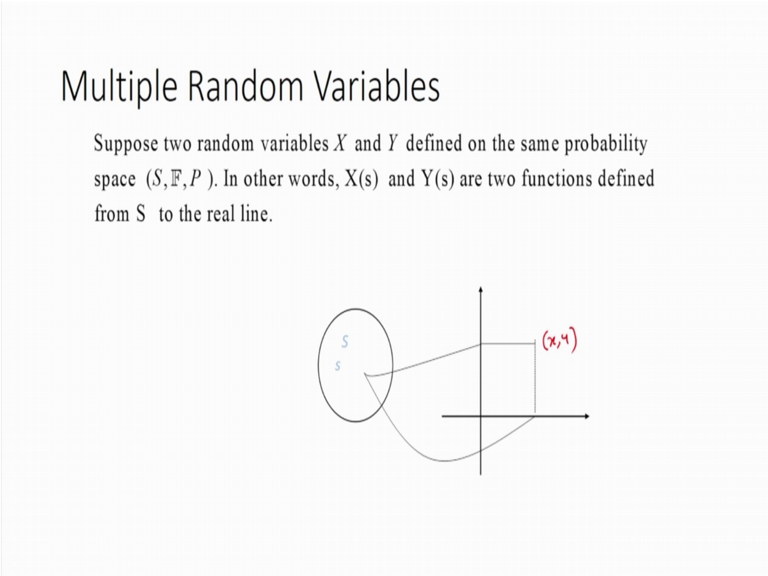



For example of a PDF it is a very famous PDF everybody knows it this is the normal or Gaussian random variable with parameter mu and sigma square. So, these are the 2 parameters the PDF is given by f x of x this is the scaling factor 1 by root of bar 2 pi into sigma e to the power minus half of x minus mu divided by sigma whole square. So, this is a bell shape suppose if it is x this is f x of x then around suppose this is mu.

So, around that above that we have a bell shape curve like this. So, this is the normal distribution above point mu this is also we will show that this is an important parameter called mean parameter and how this spread of this PDF curve will determine or control by this parameter sigma. So, that way we define random variable single random variable we also define the discrete and continuous random variables, and CDF, PDF, PMF etcetera these are the functions which characterize a random variable.

Next we will go to multiple random variable.

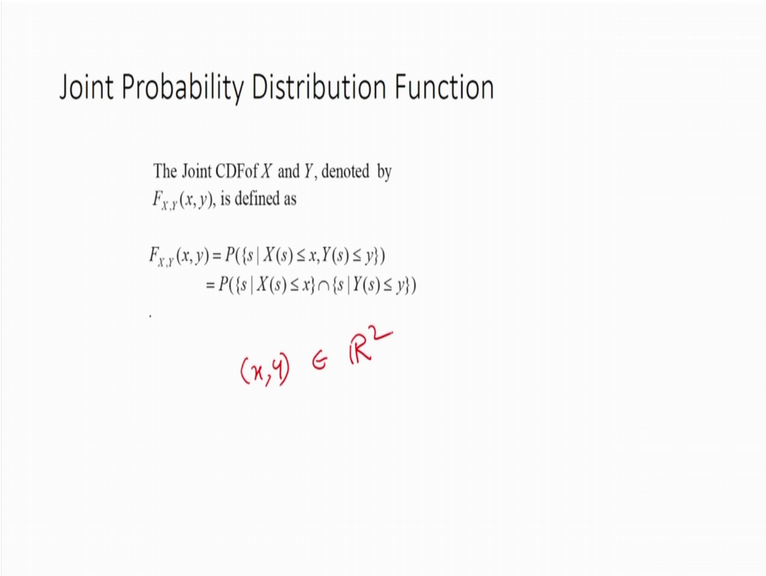
(Refer Slide Time: 15:16)




Suppose two random variable X and Y define on the same probability space structure, F, P in other words X s and Y s are two function define from S to the real line. So, we have a so at every point s we have one function X s another function Y s. So, that we can characterize these two points this point suppose which is mapped from the sample space Y the mapping X s and Y s. So, all the points can be expressed at the contagion product of these two random variables so that they can be represented in a plain like this.

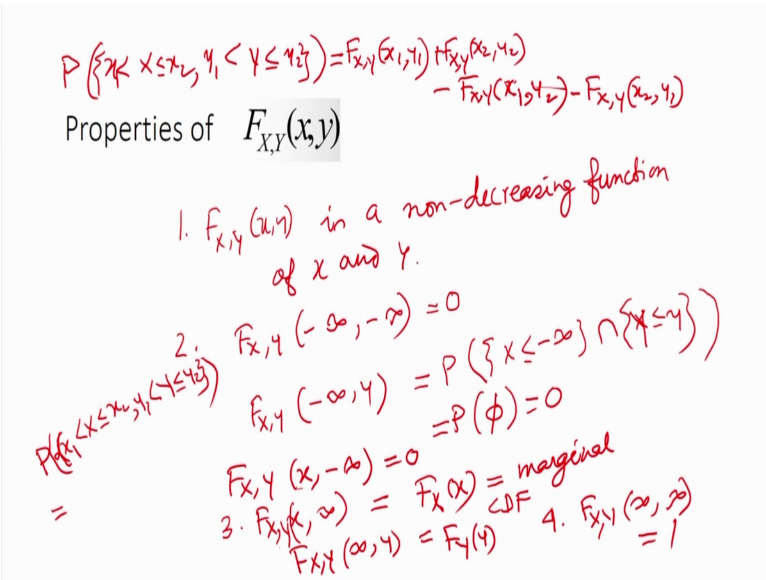
So, corresponding to this point we will have a point here like this x y of course, while defining these two random variables we take care that the probability can be properly assigned. Now given the definition of two multiple random variables we can characterize multiple random variables to some probability functions just like in the case of single random variables CDF, and PDF and PMF here also we will define the same.

(Refer Slide Time: 16:37)




Joint probability distribution function the joint CDF of X and Y denoted by this is the notation capital F subscript x y at point small x small y it is defined at the probability it is again a probability of the event s consisting of sample element sample point s such that X s is less than equal to small x and this comma means and Y s is less than equal to small y. So, this is the that is this is the probability of those x for which X s is less than equal to small x and Y s is less than equal to small y. So, this is the definition with this definition we can define what is known as the joint CDF joint CDF, and this joint CDF characterize the random variable at any point x y belonging to the plain R 2 ok.

(Refer Slide Time: 17:52)



Now, what are the properties of this joint CDF F x of y. So, they are derived from the property of this CDF basically therefore, we can say that F x of y it has two arguments X and Y it is, so joint CDF F X Y F point x y has some properties which are derived from the definition of CDF. So, these are similar to the properties of the CDF of a single random variable that way this F x of x y is a non decreasing function monotonically increasing function non-decreasing function of both the arguments of x and y.

Number 2 now let us see the event F x y at minus infinity minus infinity that will be always equal to 0 from the definition we can say that this is equal to 0. Similarly F x y at point any point you consider any point y and that is minus infinity y, so this is the probability that we can say that x is less than minus infinity that is less than equal to minus infinity that is always equal to 0 and this is and axis less than equal y is less than equal to small y. So, if we consider the intersection of these two events we will get the non event so that way probability of phi that is equal to 0.

Similarly, we can also show that F x y at any point x minus infinity that is also equal to 0, so that way F x y if any of the argument is minus infinity then it will become 0. Similarly if we consider that is important number 3 suppose what is F x of x and y argument is infinity. So, this is because infinity means this second argument that is that will correspond to the entire sample space. So, intersection of the sample space and x less than equal to small x, so that way it will be simply equal to F x of x.

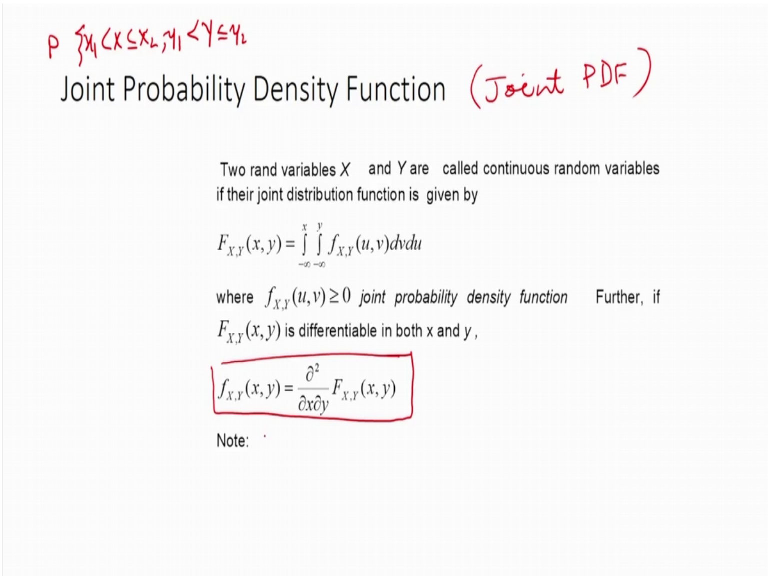
So, that way this is the distribution of small x the distribution of the single random variable at point x we can derive from the joint pd joint CDF and this is known as the marginal CDF as we derive it from the joint CDF it is known as the marginal CDF. So, the marginal CDF F x of x is obtained from the joint CDF at point x infinity.

Similarly, we can find F y of y also that is F x y at point infinity y will be equal to F F y of y. So, that way marginal CDF at point y we can obtain from the joint CDF at point infinity y. So, this is important property so that way joint CDF completely characterize two random variable X and Y and from this joint characteristics we can find out how this random variables independent individually behave that is marginal CDF in terms of the joint CDF.

Similarly, finally, we can find out number 4 this F xy at point infinity infinity that will be always equal to 1. So, these are some of the properties and we can find out any probability involving the joint CDF we can find out any probability involving the joint random variable x and y. So, how we can find out let us see.

So, we can find out for example, probability that X lies between small x 1 and x 2, Y lies between small y 1 and y 2, so this probability this type of probability we are interested to find ok. So, using the again from the basic definition we can show that this probability will be equal to probability that x lies between x 1 and x 2, y lies between y 1 and y 2. So, this is given by so this is equal to F x y at point x 1, y 1 plus F x y at point x 2 y 2 minus F x y at point x 1 y y 2 minus F x y at point x 2 y 1. So, this way we can find out the probability involving two random variable x and y.

(Refer Slide Time: 24:15)

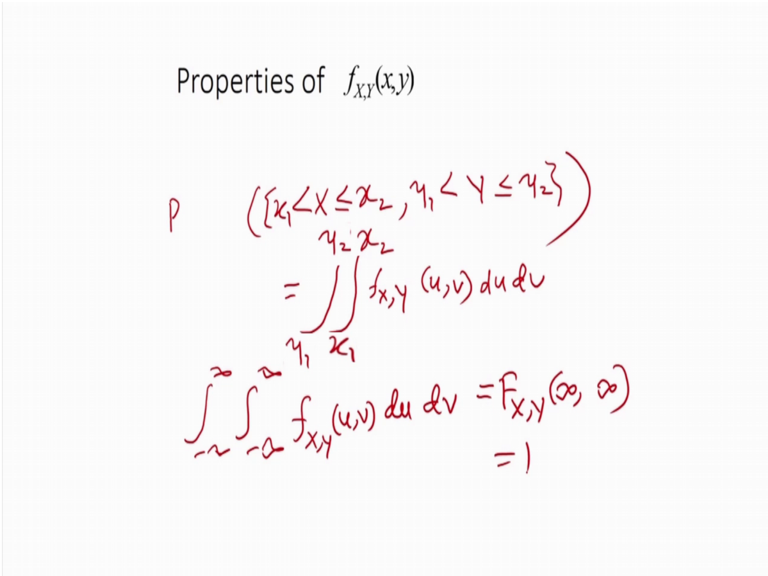



Next we will define the joint PDF two random variables X and Y are called continuous random variables if their joint distribution function is given by this. So, just like in the case of single random variable here the joint CDF is represented in terms of the double integral of a non-negative function this is the f small f x y small f x y at point u v du dv ok. Now this quantity is always greater than equal to 0 and it is called the joint probability density function joint PDF, so, that way it is joint PDF, joint PDF PDF..

Now if the CDF joint CDF is differentiable it is a partial differentiation roughly differentiation then roughly differentiable in both x and y then we can write joint PDF at point x y is the second or partial derivative del 2 del x del y of the CDF F x y at point x y. So, the second order partial derivative gives the joint PDF. Now you note that this joint

PDF is an nonnegative function it is always greater than equal to 0, that flows from the property of the joint CDF it is non decreasing both argument x and y also we can find out any probability.

(Refer Slide Time: 26:44)

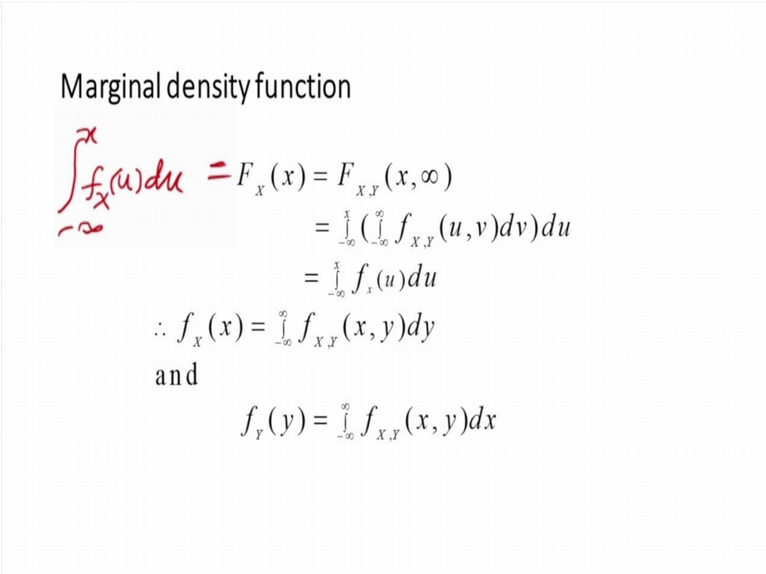



So, we can find out any probability for example, probability that x lies between some x 1 and x 2 y lies between some y 1 and y 2. So, this probability we can find out by the integration double integration of the joint CDF d v, so this is integration. So, that way we can find out the any probability involving two random variable x y, x and y two random variable x and y using this double integration.

Similarly, also if we consider suppose this integration from minus infinity to infinity minus infinity to infinity of the joint PDF du dv. So, that will be nothing, but the joint CDF at point infinity to infinity that is equal to 1. Therefore, the area under the joint CDF it will be a surface, so that way the volume under the joint CDF surface will be always equal to 1. So, that way we have defined joint CDF and joint PDF.

We will introduce also another concept what is known the independence of 2 random variable. So, we are discussing about the properties of joint PDF one more important property we will tell that is the marginal PDF how to derive from the joint PDF?

(Refer Slide Time: 28:53)



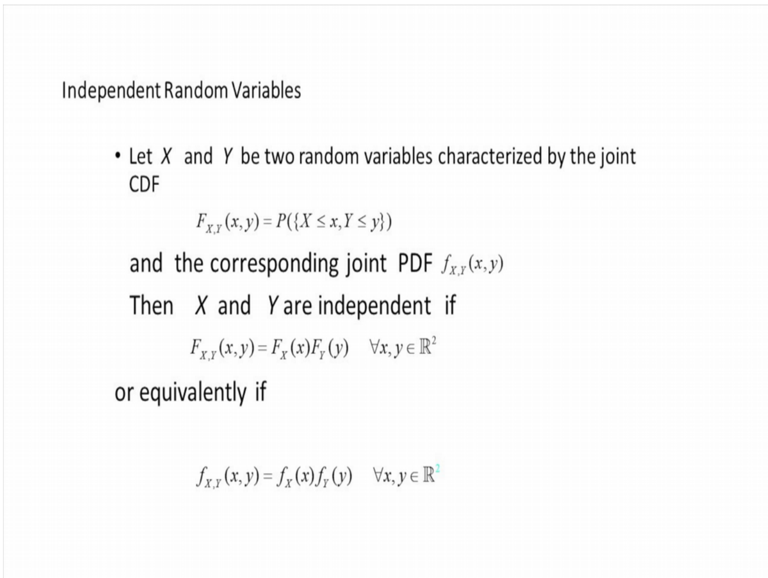

So, we know that F x of x is equal to F x y at point x infinity. So, this we can write that this is minus integration from minus infinity to x now this is infinity. So, that way F x y at point x y d y d y is from integration is from minus infinity to plus infinity and then integration with respect to x d x that is from minus infinity to x.

So, now this quantity this inside quantity is the PDF; PDF of x we can find out the marginal PDF F x of x from the joint PDF. Consider F x of x this CDF; CDF marginal CDF of x point x that is given by I know in terms of the joint CDF it is F x y at point x infinity also this quantity is equal to integration minus infinity to x marginal PDF f x u d u.

So, this quantity is same as this now this quantity joint CDF at point x infinity can be written by this integration. So, inside integration is from minus infinity to infinity with respect to y and outside integration is minus infinity to x with respect to x. So, these we can write in terms of now comparing with this is this quantity is same as F x of x. Therefore, what we get is that small f x of x that is marginal PDF is equal to the integration of the joint PDF with respect to y.

So, similarly f y of y also we can find out that is the integration of the joint PDF with respect to x. So, that way this is what is important is that given the joint PDF we can find out the marginal PDF’s both marginal PDF’s. So, that way; that means, once we know the joint PDF at each point x y we know all the characteristics of the random variable including their individual description. So, that is the importance of marginal PDF and CDF in terms of the joint PDF's and joint CDF's.

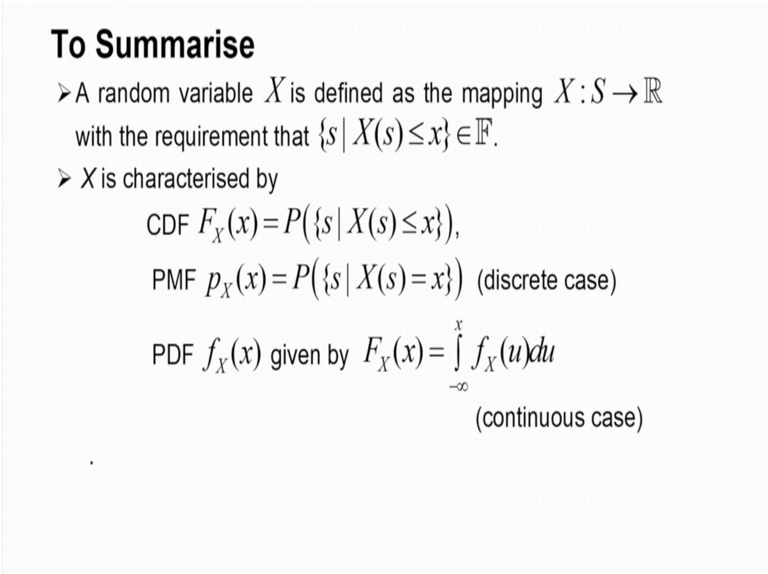
(Refer Slide Time: 31:53)




Next we will introduce the concept of independent random variables. We know two events two events are independent that is two events are independent if their joint probability is the product of individual probability. So, same concept is used here suppose x and y are two random variable characterized by two suppose x and y are two random variables characterized by their joint CDF. Then X and Y are independent it is the definition if joint CDF is the product of the marginal CDF for all x y belonging to R 2.

So, this result should be satisfied for any X Y belonging to R 2. So, we have to consider all X Y point where this relationship is satisfied and otherwise it will be called x and y will be called dependent. And in terms of the PDF joint PDF we can write that joint PDF is the product of the marginal PDF F x of x and F x of F y of y. Similarly here joint CDF is the product of the marginal CDF capital F x of x into capital F y of y. So, this is the definition of independent random variables and concept of independence plays important role in probability analysis.

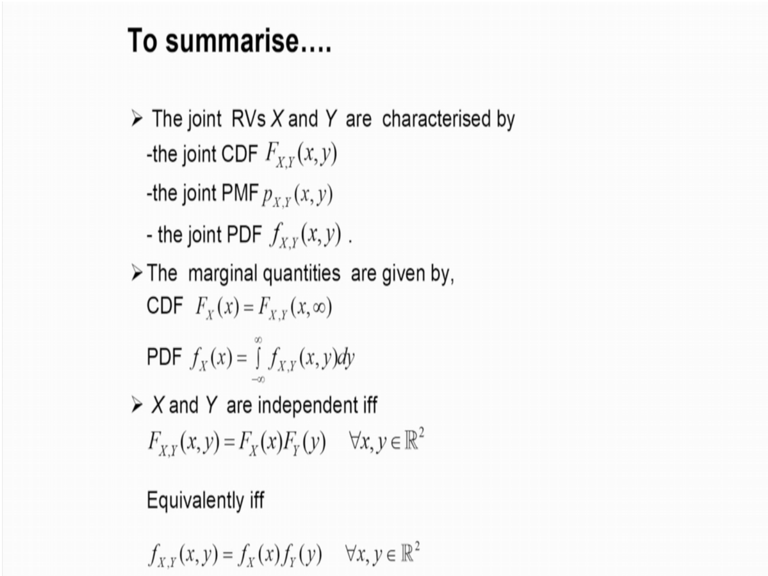
(Refer Slide Time: 33:37)




Let us summarise the lecture a random variable X is defined as the mapping from sample space to the real line with the requirement that the event s such that X s is less than equal to small x this event must belong to F. So, that way we define the random variable and a random variable X is generally characterized by the CDF that is that is the probability distribution function, or cumulative distribution function F x of x this is the probability of the event s such that X s is less than equal to small x.

Similarly, we define the probability mass function for the discrete case in that case this is the probability that X s equal to x. So, this probability is known as the probability mass function. For the continuous case we define the PDF, probability density function and this probability density function is related to probability distribution function CDF by this integration. So, F x of s is equal to integration of F x small f x u du from minus infinity to x, so that way we define the PDF.

(Refer Slide Time: 35:06)




We also discuss about joint random variables two random variable X and Y they are characterized by the joint CDF, and joint CDF can be defined similar to the CDF single random variable. We also discussed about joint PMF p x y at point small x small y this is the probability that X s taking x and Y s taking y. And similarly joint PDF that is related to joint CDF to the integration. So, this quantity we discussed.

We also discussed about the marginal quantities the given the joint CDF or joint PDF we can find out the marginal CDF. For example, CDF F x of x is equal to the joint CDF at point x infinity. So, when we put y is equal to infinity then we will get the F x of x joint CDF of x. Similarly PDF of x that is small f x of x is given by the integration of the joint PDF from minus infinity to infinity. So, this is the integration with respect to y. So, f x of x marginal PDF is given by the integration from minus infinity to infinity joint PDF integration with respect to d y.

Then we introduce the concept of independence X and Y independent if a joint CDF is product of marginal CDF that is for all x y belonging to R 2. Similarly this independence can be defined in terms of joint PDF. So, if joint PDF is product of the marginal PDF for all x y belonging to R 2, then also X and Y are independent.

Thank you.