Digital Communication

Prof. Bikash Kumar Dey
Department of Electrical Engineering
Indian Institute of Technology, Bombay

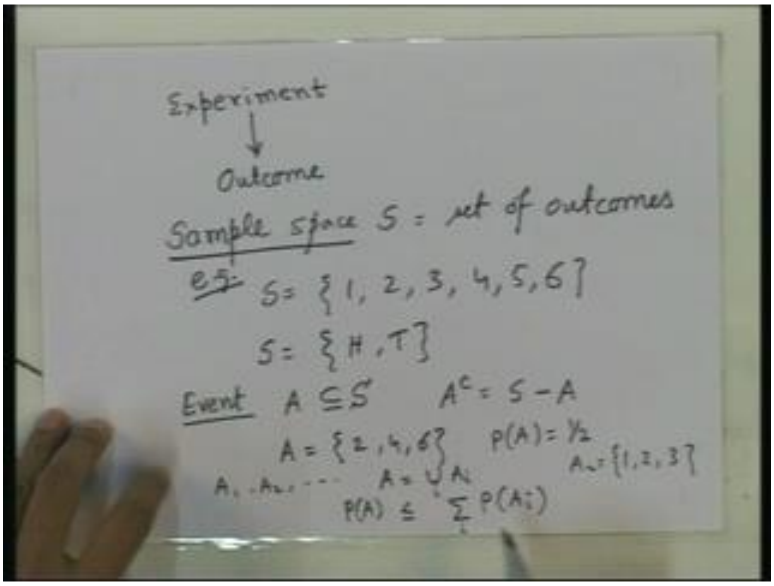
Lecture - 04
Probability and
Random Processes
(Part –1)

Hello everyone. In this class we will start with some basic background on random process probability on random process that we will need in this course. So, is most likely that this is already known to you, but nonetheless it will be used in this course. So, we will just revise the background that we need to do. So, we will start from the basic definitions of related to probability and random process. First of all the basic in the basic frame work there is a there is an underlying experiment, there may be a practical experiment or something some experiment that is performed by the nature not by anyone.

So, the experiment may be like tossing a coin, or throwing a die, or it may simply be some electrons moving inside the inside a resistor and then, the outcome there may be that some voltage observed at the ends of the resistor. In the case of throwing it die the outcome is the number that comes on the face, and the in the case of tossing a coin the outcome is head or tail. So, there is an experiment and there is an outcome of the experiment. Now, the set of outcomes is called the sample space of the experiment. So, in the case of tossing a coin the sample space is the set of the head and tail. In the case of throwing a die the sample space is the set of numbers from 1 to 6.

In the case of the voltage at the output at the ends of the resistor; the noise voltage that there resistor is generating the outcome is simply the voltage observed at the ends of the resistor. So, that voltage is continuous value it can be anything from minus infinity to infinity where as in the case of throwing a die the outcome is bound to be from one of the six numbers that is 1 to 6. In the case of tossing a coin the outcome is one of the 2 that is head and tail. So, the set of outcomes that is the sample space may be a finite set or an infinite set it may be a discrete set it may be a continuous set it can be of different types.

(Refer Slide Time: 04:00)



So, we know what is. So, there is an experiment and that is an outcome of the experiment. And, the sample space we will be denoting it by S is a set of outcomes. For example, this may be the set of numbers from 1, 2, 3, 4, 5 and 6. When the experiment is throwing a die or it may be the set head and tail. An event so, this is sample space an event A simply A subset of S. In general the definition is little more complicated there are some conditions to be satisfied, but we will not go into the details, and we will simply we will assume that the sub sets that we have considered are nice enough. And, let us simply say that an event is A subset of the sample space.

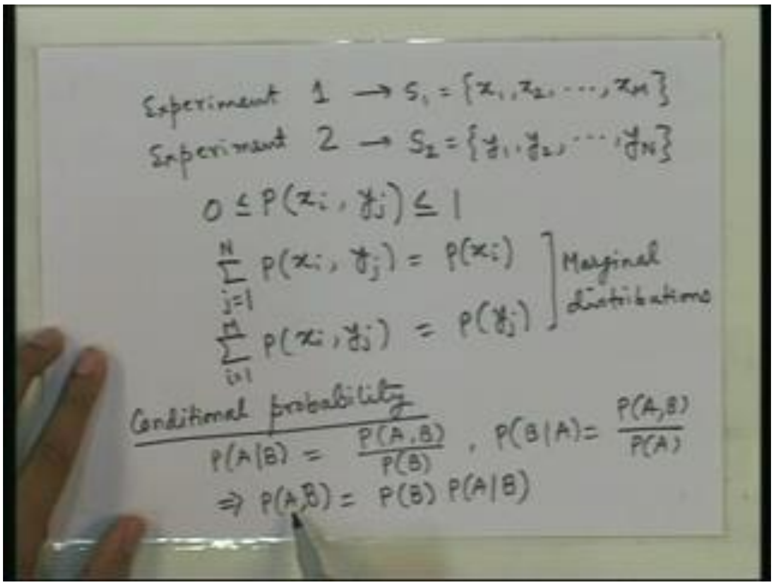
For any event A there is of course, the complimentary event that is A compliment which is set of all the set S minus the set A. Now, corresponding to an experiment there is a sample space, but at the same time every if the sample space is finite. For example, the every point in the sample has a probability of coming as outcome. So, there is a probability major associated with the sample space. As you deal with the experiment the major is on the sample space. So, for every event also there will be a probability that is the probability; that the outcome will be in that set in the set A so, that we will call the probability of the event.

For example, we can take A to be the set of even number 2, 4, 6 that is subset of this sample space and if the die is fair that is if the probability of each of these numbers is one sixth same and equal to one sixth then, what is the probability of this event A? This is half. Probability that we will observe A that is the outcome of the die will be throwing the die will be even number. So, that probability is in this case half. So, there is a probability of every element in the sample space in this case, from the sample space is finite.

Now, if we have some events A 1, A 2 and so on, some many events then we can of course, take the union of these events. Then, let us call it A that is the union of all these events and then, what is A it is basically the set of all numbers that are appearing in all these. So, the probability of A can be bounded by the summation of the probabilities it may not be same as the sum of the probabilities, but it may be even less than the sum of the probabilities. Why? Because, they are may some common elements between the events. For example A A in the we have taken one event like this. They are may be another event we can take as A 2 as 1, 2, 3 and there is this 2 is common.

So, probability of this is half probability of this event is also half, but the probability of this union this is not 1, it is not half plus half because, the probability is 1, 2 probability of 1, 2, 4, 6. So, 5 by 6 is the probability of this union this, but it will be always less than equal to the sum of the probabilities.

(Refer Slide Time: 09:14)



Now, if we have two experiments; say experiment 1 and experiment 2 with two sample spaces say S 1 which is this. For simplicity we are considering now, only sample spaces of finite size. Let us consider these two experiments with sample spaces S 1 and S 2. Then if we take any one element from here and one element from here, what is the probability? That probability of that is take x i from here and take y j from here. Then what is the probability that the outcome will be of the first experiment will be x i, and the outcome of the second experiment will be y j. That is we performed both the experiments and, what is the probability that the outcomes are x i and y j. So, this will be; obviously, in this range and this is called the joint probability major of these two outcomes.

And, once we know this joint probability distribution this P; we can find out the probability distribution of this outcome and that is called the. So, for that we have to take say for example, if you want to compute the probability of x i. So, that is the probability that the first outcome is x i and the second outcome is any of this. So, second outcome can be y 1 it can be y 2 or it can be y n.

So, we have to take the sum of all such probabilities; that is probability that the first outcome is x i, but second outcome is y 1 then probability that the first outcome is x i, but second outcome is y 2 and so on, all those. So, we have to take the sum over all j and this will be the probability of the first outcome being x i. So, if you know the joint probability distribution we can find the individual probability distributions, and these there are called them. So, if you want the distribution of this then, we have to take the summation over that is instead and this we get probability of y j. So, these 2 are called the marginal's marginal probability distributions. And, again let us consider that there are 2 events either of the same experiment, or 2 different experiments. And, we know that one event has occurred then, what is the probability the second event is second event also occurs.

So, for example, we throw one die and, we know that the event of even number has occurred. That is the outcome is an even number then, what is the probability that the outcome is 2. Then the probability is; obviously one-third, if the die is fair because if you know that the event is even the outcome is even. Then, there are 3 such numbers with equal probability each of them having probability then one-third.

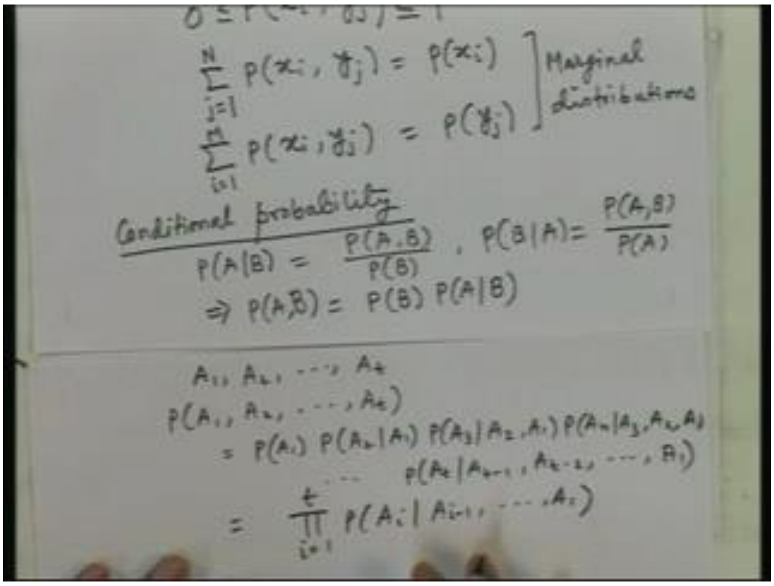
So, that to deal with that kind of situation we need to define what is called as conditional probability. So, if there are 2 events probability of this A and B then, if B is known to occur B has occurred then what is the probability that A is also true that A also occurs.

So, that is defined as A B by the probability of B. So, this is the marginal distribution this is the probability that both A and B is true.

So, these events may be of two different experiments or may be the same experiment. Similarly, probability that B given A is probability A B by probability of A. And from these 2 one can see that from here we can write probability of A B is probability of B times probability of A given B. So, this is quite expected because, what is this is the probability that both A and B are true, both the events A and B occur. Then, that is same as the probability that B occurs and then given that B occurs what is the probability that A also occurs.

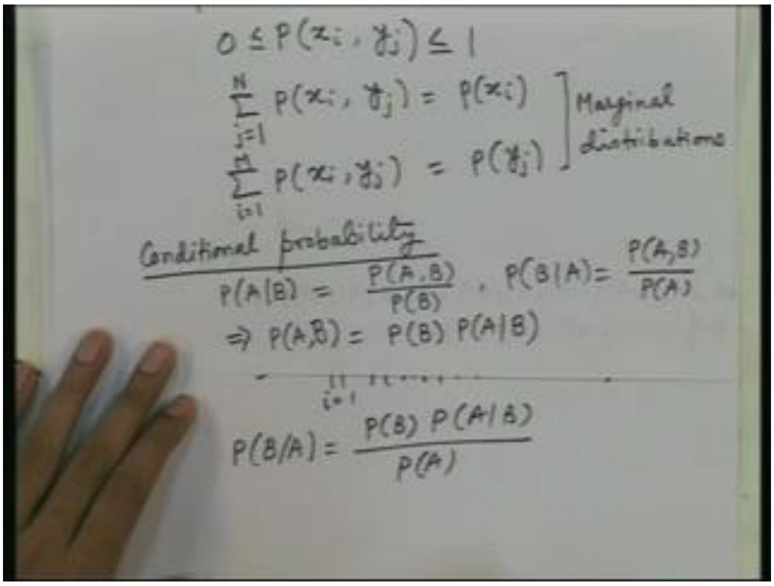
So, that if you multiply those two probabilities that is what is the probability that B will occur and then, what is the probability that if B occurs, what is the probability that A will also occur. So, those 2 probabilities together will give you the probability that A and B both occur. So, this is also called chain rule of probability, so will see a generalization of that.

(Refer Slide Time: 15:59)



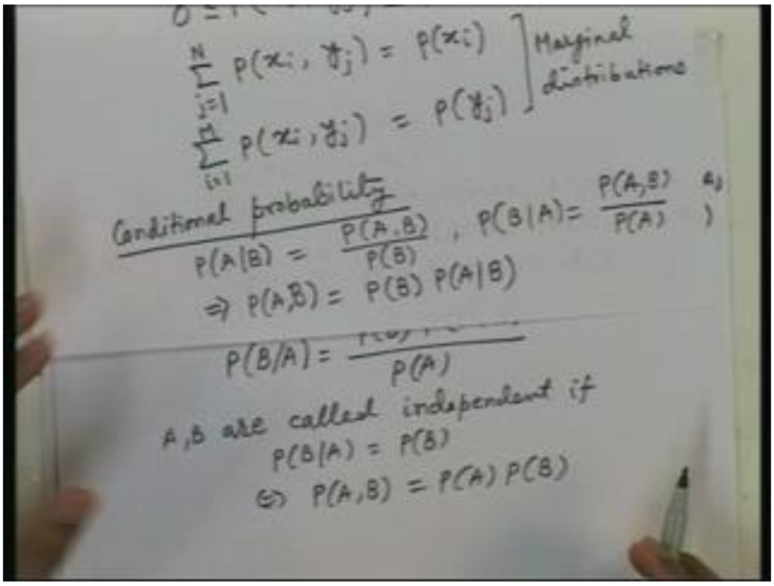
The generalization is; if you have now, T events A 1, A 2, A t one can show that this probability is probability of A 1 time probability of A 2 given A 1, probability that A 3 given A 2 A 1, probability of A 4 given A 3 A 2 A 1 and so on. Till probability of A t given A t minus 1, A t minus 2 and so on, till A 1. This can be written in short as probability A i given A i minus A to 1. So, this is called chain rule. And, now from here one can say that one can now express suppose we know that; we know this probability, but we want to compute this probability. So, what is the way to compute B given A probability of B given A from probability A given B. So, in this expression this does not appear, but then this one can be expressed in terms of this here is how it can expressed interest of this. So, we can write probability B given A as probability of A B that is this, probability of B times probability of A given B by probability of A.

(Refer Slide Time: 17:50)



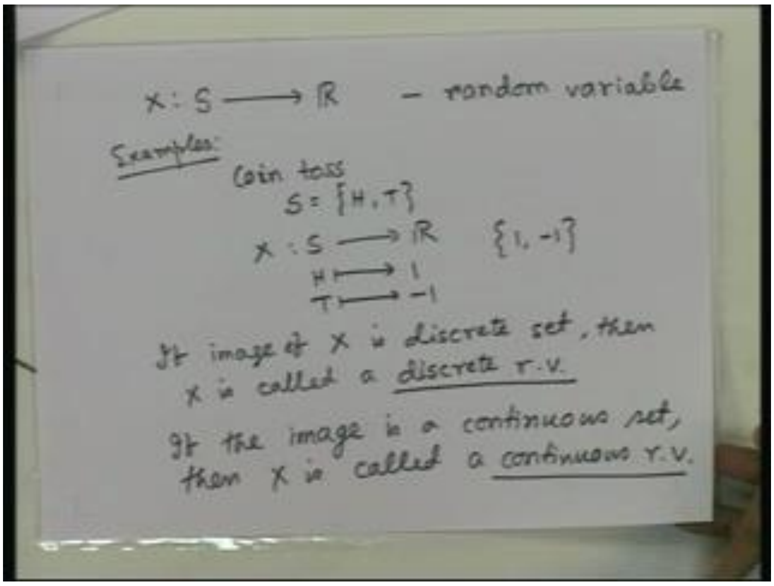
Now if 2 events of there A and B they are called independent, if this probability is same as. So, A and B are called independent events, if P B given A is P B.

(Refer Slide Time: 18:18)



So, if we simply put this quantity equal to P B what do you get this P B P B cancels from both sides and then we will get P A given B equal to P A. So, that is what is also that is equivalent. Now, on the other hand if we put here this equal to P B what do you get we get P A given B equal to P A times P B. So, we get this is same as saying that P (A, B) is P A times P B. That is the probability of A and B together is equal to probability of A times probability of B. So, this is also same as saying P A given B is P A. So, these are all same equivalent and, if this is true then this to this 2 events A and B are called independent events.

(Refer Slide Time: 20:12)



Now, we have a sample space S cross one into an experiment, and if you considered a mapping x of S into R. That is what is this mapping? This takes every element from the sample space to a real number. So, the outcome of the experiment may not be a real number for example, the tossing a coin the outcome either head or tail it is not a real number. So, but we can associate with the every outcome a real number. So, that for every experiment the outcome is mapped to the set of real numbers. The mapping can be indifferent ways, but we can map to the set of real numbers, and such a mapping is called a random variable.

So, there is an experiment and there is a mapping of the A set of outcomes to the set of real numbers and that mapping is called the it called a random variable. So, if you have a random variable like this. So, this is a random variable. For the let us see some examples.

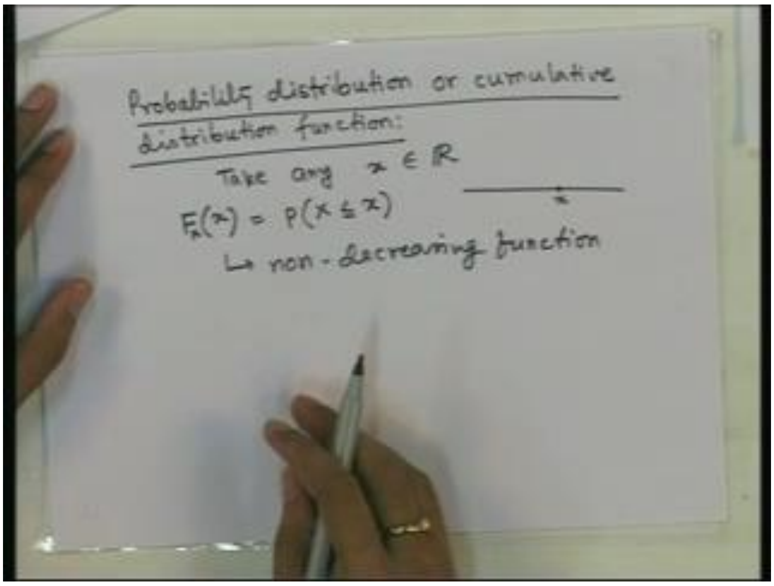
So, consider coin toss. So, S is head and tail. Now, we can consider a mapping like this where head is mapped to 1 and tail is map to minus 1 or tail is mapped to 0 that will be a different random variable. So, if we can take different mappings and they will be different random variables. So, this is an example of the of a random variable.

Now, if the outcome if the if the image of this mapping is a discrete set it is not continuous value. Then the. So, if image of X is discrete set then X is called a discrete random variable. So, X is called a discrete random variable. For example in this case it is the discrete random variable the mapping because a image of the mapping is 1, minus 1 it has only 2 elements and this is a discrete set. On the other hand if you take the voltage across a resistor the lowest voltage that thermal lowest that is generator of the at a resistor it is continuous value. That is also a random variable the outcome itself is a real number and the mapping can be taken as just the real number itself the identity mapping.

So, every real number goes to itself that is the mapping then that is the image of the random variable the mapping is a continuous set it can be anything it is not either plus 1 minus 1 or plus 2 0 it is not integer, but it can be anything. So, it is a continuous random variable.

So, if the image is a continuous set then X is called a continuous random variable. Now, let us define some more probabilistic functions and concepts regarding random variable.
So, the random variable the values taken by random variable are real numbers unlike the outcome of an experiment. Which may not be a real number like head and tail for coin toss they are not real numbers, but the random variable is defined in such a way that the out the value is a real number. So, we can now talk about what is called a probability distribution or also known as cumulative distribution function.

(Refer Slide Time: 24:43)



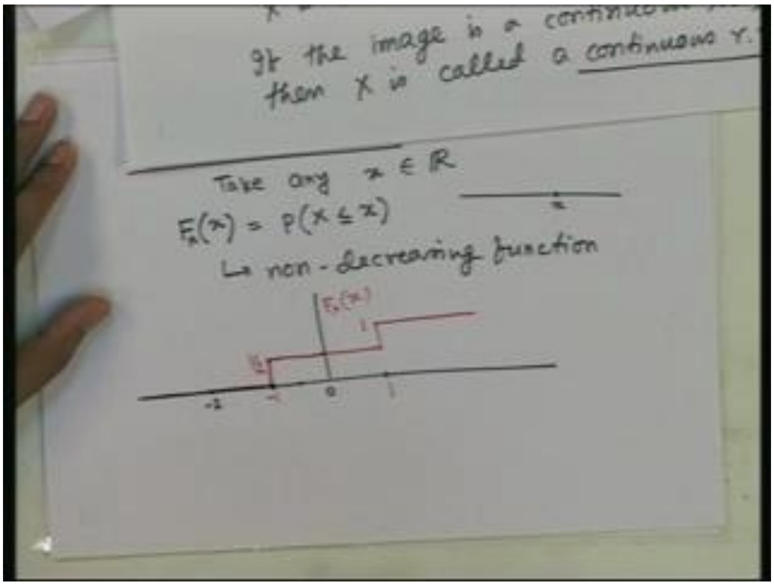
So, it is basically for any X take any X which is real number. So, we denote the set of real numbers by this symbol that is real line. Then take any real number X which belongs to this set. Then the probability distribution function of the random variable X. So, we have a random variable X. So, for that random variable this function Fx is defined in the following way. What is Fx at x; x is x denotes a value a real number. So, where as this X is the random variable the mapping. So, X this Fx at X is basically the probability that X is less than equal to x. So, that is if you take a value x what is the value of this function at x the value is the probability that the random variable capital X will be less than x less than this. The probability that the value of the random variable will be on the left side of x. So, that is the that probability is the value of the function at x.

So; obviously, that is less than one, but greater than 0 it is between 0 and 1 because it is a probability and also it is increasing as x increases because if you take one x equal to 1 the value is the probability that the random variable is less than 1. Now, if you take the value 2 the value of the function at 2 is the probability that the random variable is less than 2 which will be more than the value at 1. So, it is increasing as x increases. So, this is increasing or whether non decreasing it does not decrease. So, it is non decreasing function. Let is see some example. So, we just now considered this random variable head tail mapped to 1 and minus 1. Now, if let us considered a fair coin where the probability of head and tail are same that is the half then let us draw the let us plot the cumulative distribution function for this random variable. So, this 0.

Let us consider say minus 2. So, at minus 2 what is the value what is the value is probability that the value of the random variable is value of this random variable is less than minus 2, but the value of this random variable cannot be less than minus 2 because it takes either 1or minus 1 values.

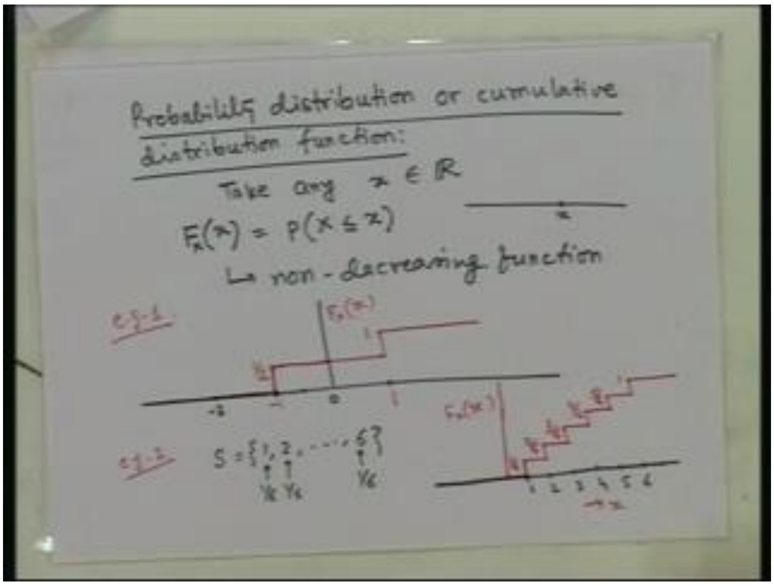
So, the probability is 0 at minus 2. So, for any value less than minus 1 the values any real number less than minus 1 the value of the distribution function will be 0. So, it is 0 here. So, it is the value is 0, here and at minus 1 what is the value it is still 0 because it the probability that it is less than equal to minus 1. Now, suddenly here the value becomes half because the probability that the random variable is less than equal to minus 1 is half. It is equal to minus 1 with probability half. So, it the value is half here. So, this is half. And then probability that here the value is on the left hand side of this is also half because it takes value already this these values are not taken. And this probability is half. So, it is still half it does not increase it is still half and it increases again at 1 because this is the next value that the random variables takes. And, here at 1 what is the value it is the probability that random variables is less than equal to 1 and that is 1 the probability is 1 because it is always less than equal to 1. The value of the this random variable is always less than equal to 1. So, it jumps to one here.

(Refer Slide Time: 30:19)



So, this value is 1 and there it remains at 1 it cannot increase more than 1 of course, because this is probability it cannot be more than 1. So, this is the cumulative distribution function of that random variable. Now, if we consider the this is example 1; for coin toss and second example we throw a die.

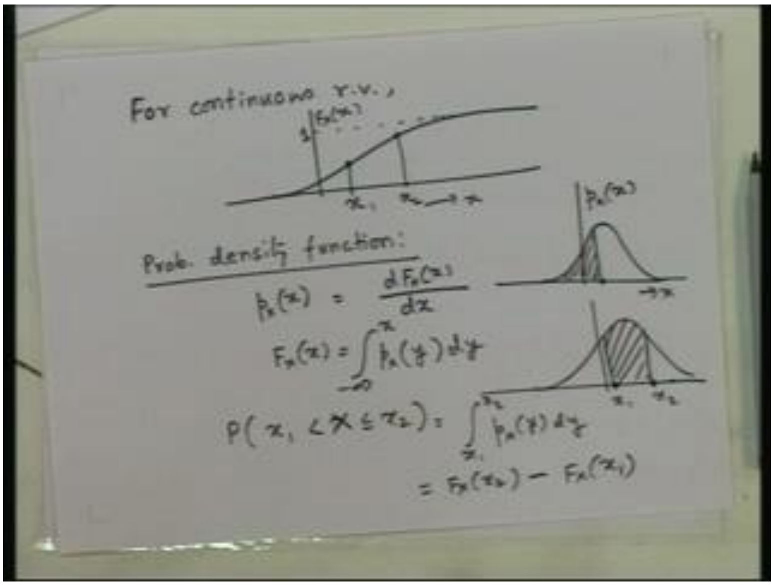
(Refer Slide Time: 30:48)



S is 1, 2 to 6 and each has. So, it is fair die. So, each has probability 1-6. Now, so, what is the cumulative distribution function of this of this random variable let us take 1, 2, 3, 4, 5, 6. So, if you take a number below one the random variable cannot be less than that. So, probability is 0 everywhere here, the probability is 0 and at 1 the value is 1 minus 6. So, it rises to 1-6 and then at 2 it rises to 2 by 6 and at 3 it rises to 3 by 6, 4 by 6, 5 by 6 and to 1. So, this is one this is 5 by 6 this is 4 by 6. So, this is the cumulative distribution function of this random variable.

Now, for a continuous random variable it will not a cumulative distribution function. Will not look like this kind of staircase because the in the that case there the probability that the random variable takes a particular value will be 0. Because there are infinite number of values it can take. So, usually it will whereas, here there are finite number of points where it takes values there they will be infinite number of points between any range between 0 to 1 itself 3 will be infinite number of values all the values can be taken. So, if they are all more than 0 s then the total probability will be more than 1. So, it is it rises continuously instead of like this.

(Refer Slide Time: 33:40)



So, for continuous random variable a typical cumulative distribution function looks like may look like this. So, the maximum value is of course, 1 the minimum value is 0. So, for any cumulative distribution function of course, at extensive infinity that will it will go to 1 and it extends to minus infinite the value will go to 0. So, for a continuous random variable it rises smoothly and if the derivative exists then the that derivative is called the probability density function. So, if we take we define another quantity probability density function it is basically derivative by this and it is d Fx by dx. So, for this kind of cumulative distribution function if you plot the derivative of this function it will look something like something like this.

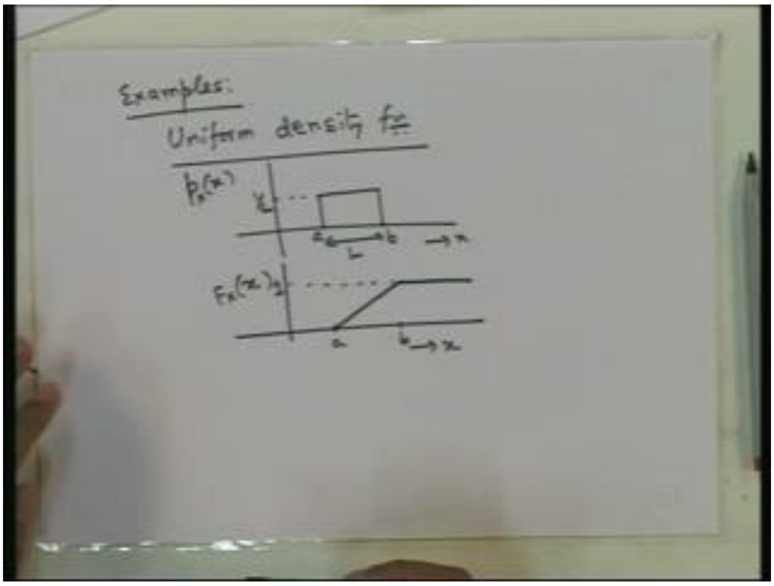
And from this definition one can see that this means F x is nothing, but the integral of this. So, we need to take an auxiliary variable to integrate d y from minus infinity to x. So, this is basically the area under this curve till x. So, if you this value this value is nothing, but take this value here this area this area is the probability that the random variable its value is less than this and that is same as this value that this what is the cumulative distribution function. So, if we know the cumulative distribution function the way to get the probability density function it to differentiate it. If you know the density function we can get the probability distribution function by taking the integral minus infinity to x plot that as function of x we will get the cumulative distribution function.

Now, if you know the density function then can we compute probability of this type probability that the random variable is between two values. Suppose you have to find out what is the probability that that say we have the density function say like this and we have taken x1 and we have taken x2. What is the probability that the random variable will be in this range the value will be in this range. That is the area under this curve from x1 to x2 because this is the density this really the density of the curve that is density of the probability that why it is probability density function. The probability that the value is in this range is simply the area under this area. So, this is integral x1 to x2 P xy dy. And we can of course, write these in terms of the cumulative distribution function also.

This is this area is nothing, but the area till x2 minus the area till x1. So, area till x2 is Fx x2 minus infinity to x2 that area is this value for this value is for this value is the area from minus infinity x2 in this curve and this value is the area from minus infinity x1 here

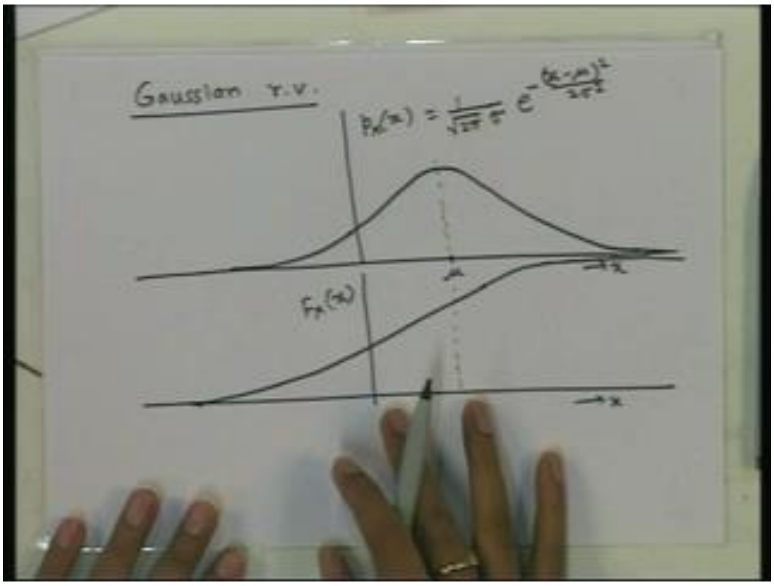
So, if we subtract that we will get this area extra area that is this minus this. So, this minus Fx x1. Now, let us see some. So, we will just summarize by saying that this is the cumulative distribution function this can be defined for any random variable where as the probability density function is defined only for continuous random variable. And let us see some examples of continuous random variable density function of continuous random variables.

(Refer Slide Time: 39:17)



Uniform density function the density function can be like this say it starts from some were and say that with this we say l and then this value will be one by l because the area under this curve the total area must be 1. The total probability is one. So, this basically says that the this is uniform the probability distributed uniformly in these 2 ranges this is in this range a to b And what will be the cumulative distribution function for this random variable the cumulative this is Px what will be Fx for this take a b it will rise linearly. So, this value is 1 and slope of this is basically this value 1 by L is slope of this curve this line the derivate of this is this.

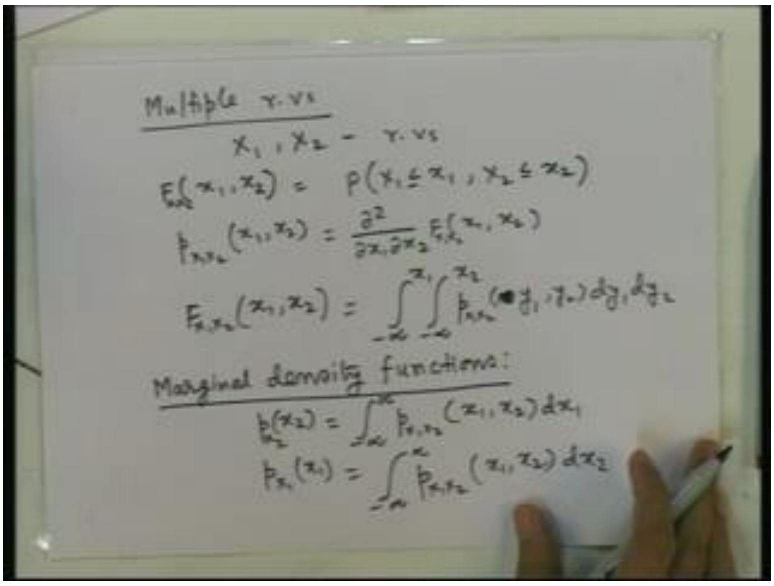
(Refer Slide Time: 40:57)



Now, let us take another example which very important in this course Gaussian random. Variable for Gaussian random variable the density function is 1 by root over 2 pi sigma into the power minus x minus mu whole square by 2 sigma square. Mu and sigma are 2 variables mu is basically what will define later as mean that is the center of the density function. So, suppose this is mu and take some sigma to be something some real number then some positive real number. Then the this will look like this going to 0, but it is never touching 0 actually. So, it is maximum at mu and symmetric around mu also. So, this the probability density function it is a the bell shaped famous bell shaped curve and the if you plot the integral from minus infinity to x as function of x what will you get it will be like this. The slope is maximum here because this value is maximum this is the slope of this curve. So, this is Fx of x there is no closed formed expression for this is the integral of this for minus infinity to x.

Now, let us consider multiple random variables. Suppose we have multiple random variables just like we considered two experiments and considered joint probability we will.

(Refer Slide Time: 43:16)



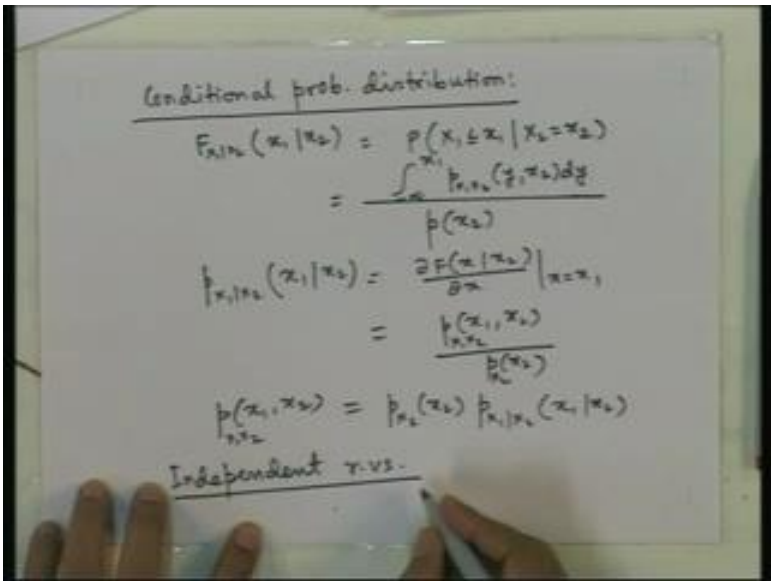
We now define joint distribution say we have x1 and x 2 are random variables and then we can define F x1 x2 the joint cumulative distribution function, as it is two dimensional function now we take two variables x1 and x2 and this is probability, that x1 is less than x1 less than equal to x1 and x2 is less than equal to x2. So, this is the function of two variables. So, to plot it you need three dimensions two dimensions for the variables and one dimension for the value. So, we can now differentiate it with respect to these two variables and get the density function P x1 x2 at x1 x2 this is del square del x1 del x2 F x1 x2 x1 x2 the derivative of this.

And then we can write this in terms of this also simply as integral this is basically minus infinity to x1 minus infinity to x2 then P. Let us call this twosome get some auxiliary variables d y1 d y2 the integral. From this density function we can get the marginal density function of these two random variables separately for x1 there is at density function for x2 also there is a density function. If they are if they are continuous random variables then this density function is the joint density function of two random variables

If you know the joint density function of two random variables this gives us all the information not only about two of two random variables, but; obviously, we can extract any individual information about any other random variables like the probability density function of the two random variables separately that is the marginal density functions. So, the way to get them is as. So, the marginal density can obtained as p x2 of x2 this is marginal density function of x2. This minus infinity to infinity p x1 x2 this is x1 x2 d x1 with respect to x1 if you integrate then you get the marginal distribution of marginal density function of x2.

Similarly, you can get marginal density function of x1 by integrating the joint density function with respect to x2. So, we have multiple random variables we have a joint cumulative distribution function and a joint density function and from the joint density function. We can get the marginal density function by integrating with respect to the other variable. We can define just like for events we can define for random variables we can define conditional probability distribution conditional density function and all those related quantities.

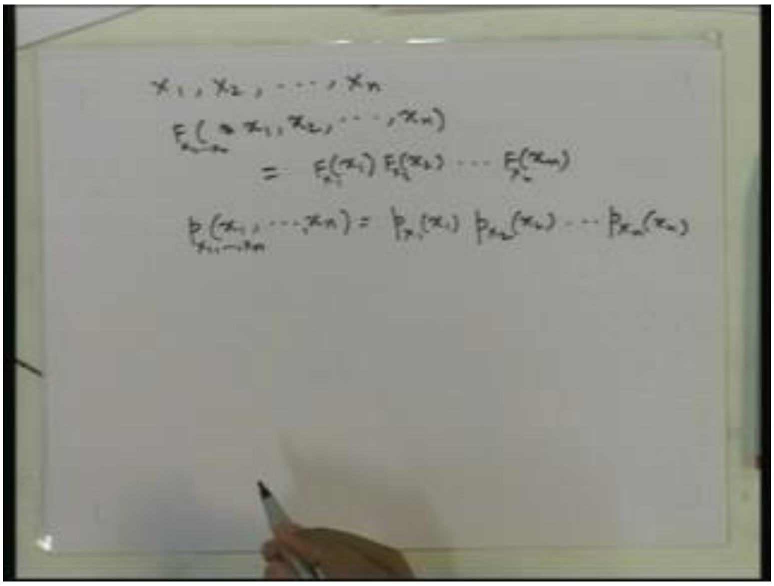
(Refer Slide Time: 47:57)



So, the conditional probability distribution if you take x1 given x2. So, x1 given x2 that is if you know that the value of x2 is x2 then what is the cumulative distribution function conditional cumulative distribution function of x1. So, that is defined to be x1 what is probability of the x1 less than x1 given that x2 is x2. So, this can be shown to be minus infinity to infinity x1 the integral joint density function this one can show that this is same as this. Now, from here one can get conditional density function at x1 given x2 this will be simply del F x given x2 by del x at x equal to x1 and this nothing, but p x1 x2 at x1 x2 by p x2. So, this is looks just like conditional probability of two events one event given the other.

This is the joint previously you have joint probability and divided by the probability of this. So, the here also similar quantity, but these are not random variables these are density function of random variables. And. So, from here we have chain rule that p x1 x2 is p x2 x2 times p x1 given x2 x1 given x2. If you have more random variables we can generalize this just like we did before and if now this is same as p x1 for all x2 and x1 then x1 and x2 are called independent random variable. So, independent if ((when at)) random variables are called independent if we have this random variables then they are called independent random variables if their cumulative distribution function factors this cumulative distribution function is the product of the marginal cumulative distribution functions.

(Refer Slide Time: 51:22)



So, this is this times F x2 times F xn. So, all the random variables are actually independent the value of one random variable does not get influenced the probability of x1 being something is not influenced by what happens to the other random variables this basically that. So, this is true and at the same time this true one can also show that this x1 the this density function is also can be factored. So, these are x1 x2 xn. So, x1 x1 p x2 x2 p xn xn.

So, we have in this class we have discussed what is an experiment what is the outcome of an experiment and then we defined the sample space to be the set of outcomes of an experiment and then we defined independent events conditional probability of one event given that other another event has occurred and then we defined random variable to be a mapping of the samples space into the set real numbers.

And, since the sample space has a probability associated because of the experiment the random variable the value of the random variable also has some probability major on the values. So, using that probability we defined what is called the probability distribution function or cumulative distribution function. And, then for a continuous random variable that cumulative distribution function is differentiable and the derivative is the called the density function and we have actually seen why it is called density function by integrating that density function in one range get we actually get the probability the random variable will in that range.

So, it makes sense to call it the density function and then we defined joint cumulative distribution function joint probability density function of some random variables multiple random variables and we defined conditional probability distribution and conditional density function and then at the end we have defined independence random variables. We will we will continue or discussion probability random variables in the next class also.

Thank you.