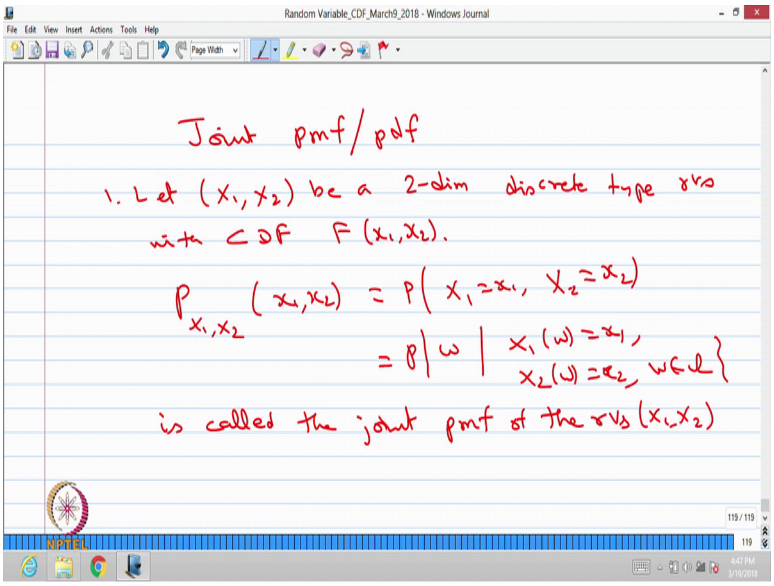
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 25

(Refer Slide Time: 00:04)



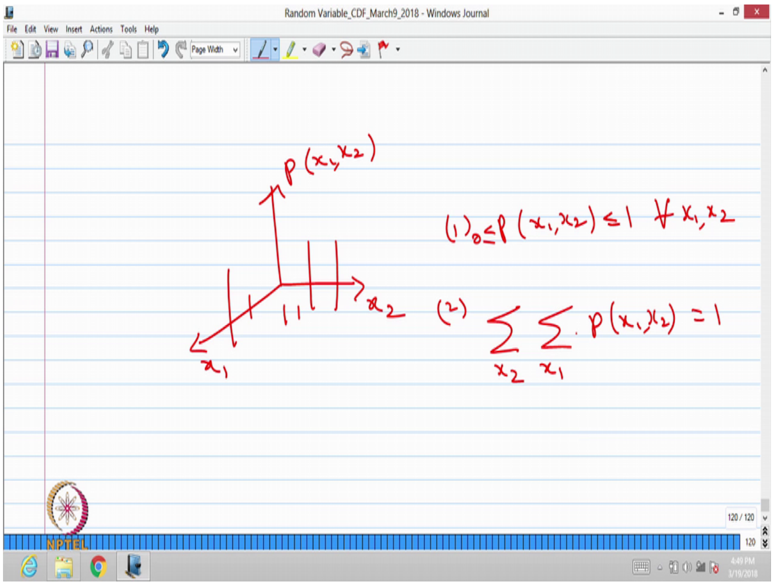
Now, we will move into joint probability mass function, and joint probability density function of n dimensional random variables, or random vector of size n.

Let me start with the joint probability mass function. Let me start with the 2 dimensional that is easy. Let X 1 comma X 2 be a 2 dimensional discrete type random variables; that means, X 1 is a discrete type random variable, as well as a X 2 is also discrete type random variable with the CDF capital F of x 1 comma x 2.

One can define the probability mass function in together; that is ah, probability of x 1 comma P of x 1 comma x 2. That is with the variable x 1 comma x 2; that means, the probability of X 1 takes a value small x 1, and X 2 takes a value small x 2. Where small x 1 is the images of X 1 or ranges of X 1, and small x 2 is the ranges of the random variable X 2 or the images of X 2. Put together that is the probability of X 1 takes a value small x 1 x 2 takes a value small x 2. This is nothing but the p of collection of w; such that X 1 of w that is equal to X 1 and X 2 of w that is equal to x 2 and w belonging to omega; that means, this is the event collection of possible outcomes satisfying this event.

So, p of this event that is the probability of event satisfying this condition. So, this function is called the joint probability mass function of the random variable X 1 comma X 2. This is the probability mass at the point X 1 and X 2.

(Refer Slide Time: 03:08)

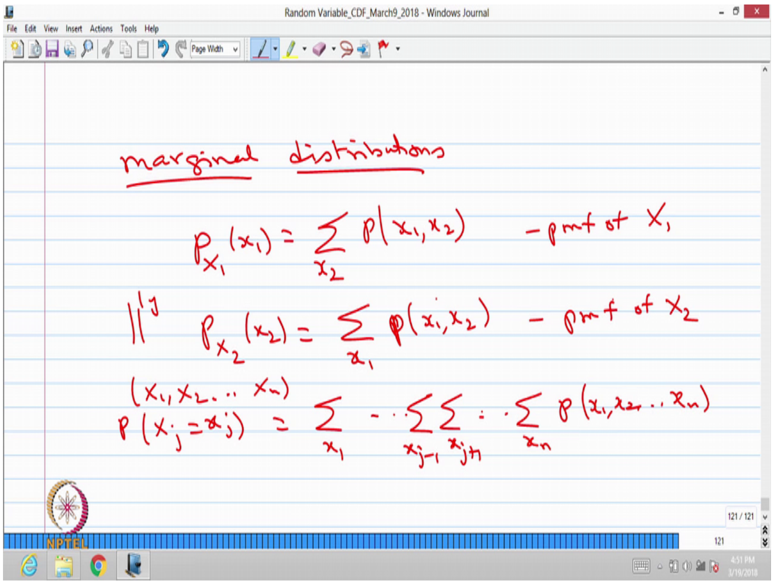


You can go for graphical representation of joint probability mass function x 1 x 2. So, this is the probability mass function of x 1 comma x 2; that means, at some point in the 2 dimensional plane x 1 x 2 plane, whatever the smaller heights whatever the heights, that is going to be the probability mass function at the point x 1 comma x 2.

Both are discrete type random variables; therefore, this can be represented in the 3 dimension plane x 1 is the one axis coordinate, and x 2 is a another coordinate, and height is z axis is the probability at the point x 1 comma x 2. The joint probability mass function satisfies 2 properties this is always going to be lies between, it is always lies between 0 to 1 for a every x 1 comma x 2.

The second condition if you make a double summation of probability mass function at the different x 1 x 2 that is going to be 1; that means, if you add all the heights over the X 1 X 2 plane that addition is going to be 1; that means, wherever there is a mass it has greater than 0 if you had all the masses that is going to be 1. From the joint probability mass function, one can get the probability mass function of x 1 and x 2, they are called marginal distributions.

(Refer Slide Time: 05:08)

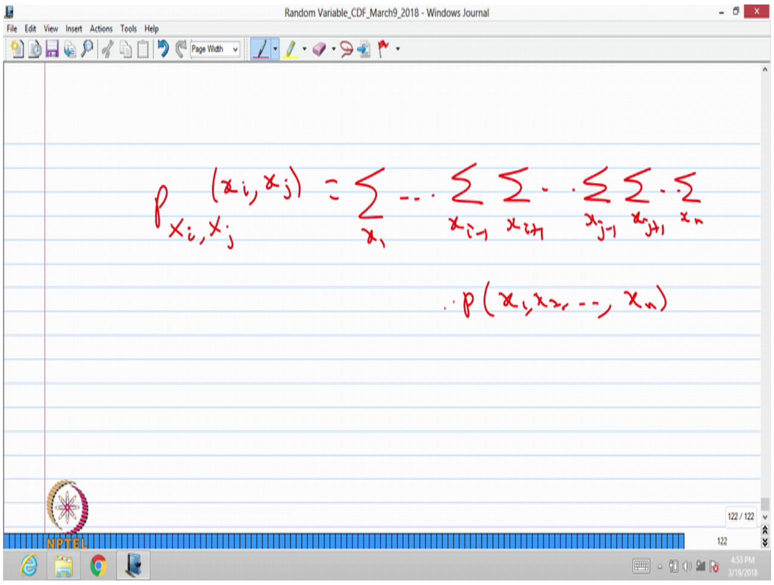


That means if I want to find out the probability mass function of x 1 from the joint probability mass function of x 2 by summing it over x 2, I can get the probability mass function of x 1. We can verify whether this is going to be the probability mass function, in this summation this values always going to be greater or equal to 0, lies between 0 to 1, and if you make a summation over x i, x 1 that is going to be double summation over x 1 and x 2. That is going to be one therefore; this is the probability mass function of the random variable X 1.

Similarly, one can find the probability mass function ofx 2 by summing over x 1 of joint probability mass function so, this is the probability mass function of X 2. The way we have done, we can go for n dimensional random variable, then we can get the probability mass function of any one random variable by summing over the joint probability mass function of x 1 to x n except j th variable x j.

We can get the marginal from the joint distribution from the joint probability mass function of n dimensional.

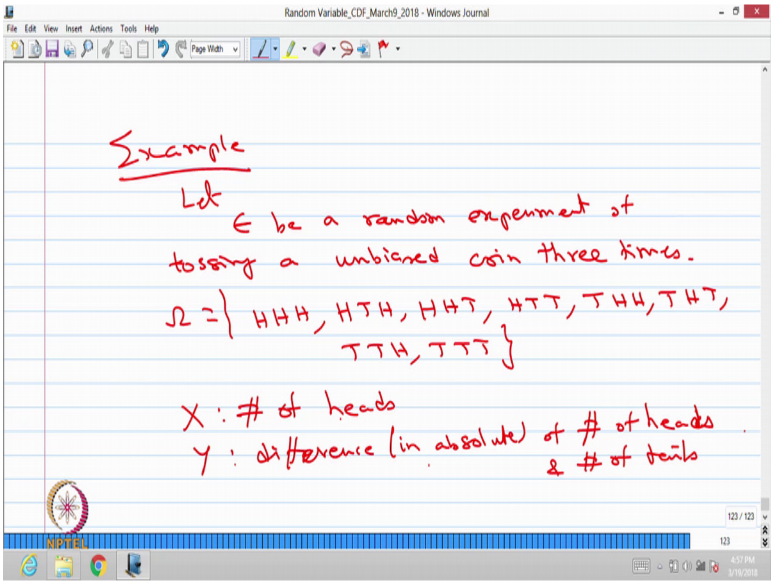
(Refer Slide Time: 07:24)



I can find the joint distribution of xi and xj by summing over x 1 to xi minus 1, xi plus 1 x j minus 1 x j plus 1 till xn of joint probability mass function of x 1 to x n; that means, by n minus 2 summations without xi and xj, one can get the joint probability mass function of xi comma xj; that means, always from n dimension random variable either CDF, or if they are discrete type random variable, you can get the lesser distributions of jointly by summing it over the other variables.

So, by doing again and again you can get the marginal distribution of one random variable. So that means, from n random variables you can get the joint distribution of n minus 1, then n minus 2 and so on finally, you can get the marginal distribution of a any random variable. Let us go for one simple example how one can visualize the 2 dimensional discrete type random variable as a example.

(Refer Slide Time: 09:22)



Let E be a random experiment of, E be a random experiment of tossing a unbiased coin 3 times. The random experiment is tossing a unbiased coin 3 times.

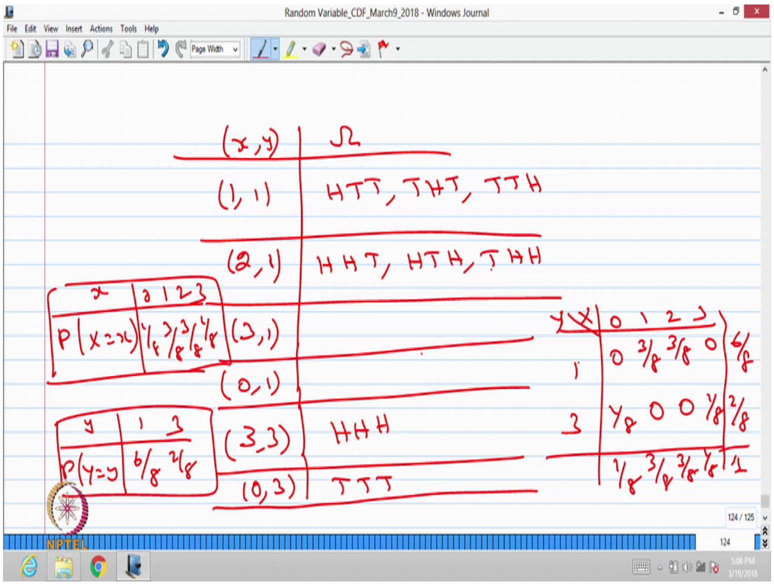
Therefore the omega is going to be the collection of all possible outcomes. That is I use the notation H for getting head, T for tail. So, since we are tossing a unbiased coin 3 times, therefore, you will have a 2 power 3 so, you have 8 possibilities. So, head head head head tail head or head head tail, and head tail tail, then tail head head, tail head tail, tail tail H, then last tail tail tail. So, these are all the 8 possibilities, or 8 possible outcomes of this random experiment of tossing a unbiased coin 3 times.

Now, I am going to define 2 random variables in this random experiment. And our interest is to find out the joint distribution of this 2 random variables, first let me define first random variable x as a number of heads in tossing a unbiased coin 3 times. The random variable y is nothing but difference in absolute of number of heads and number of tails.

You see it very carefully, the random variable axis number of x whereas, the random variable y is difference in absolute of number of heads and the number of tails. Therefore, you should know what are all the possible values of X, what are all the possible values of Y, then you can conclude what type of the random variable X and Y. Then you can go for finding out the distribution based on whether it is a discrete or continuous. The way the axis define the number of heads, and the random experiment is a tossing a coin unbiased coin 3 times. Therefore, there is a possibility you will get no times head or one times head or 2 times head or 3 times head.

Therefore the possible values of x; that is 0 1 or 2 or 3, whereas, the y is the difference in absolute of number of heads and number of tails, therefore, the possible values of y is going to be y 1 and 2 because of a difference in absolute.

(Refer Slide Time: 13:47)



Therefore you can go for make out the table of different values of x comma y, and what is the collection of possible outcomes which is going to give the values of x comma y.

For example, suppose you go for x takes a value 1 y takes a value 1; that means, number of heads is 1, and the difference in absolute with a number of heads and tails that is also 1; that means, a the possible outcomes from the omega; that is head tail tail or tail head tail or tail tail head. All these 3 possibilities gives the value of x comma y is 1 comma 1, ok. Similarly you can go for what are all the possible outcomes in which gives the values 2 comma 1, that is going to be the number of heads is going to be 2, and the difference in absolute with the number of heads and tail that is going to be 1, therefore, it is going to be head head tail, head tail head tail head head. The next one you can go for finding 3 comma 1.

If you go for 3 comma 1, you will get no possible outcomes. Similarly, if you go for 0 comma 1, there also you would not get any possible outcomes. If you go for 3 comma 3 number of heads is 3, and difference in absolute heads with a tail that is also 3; that is possible with the head head head. Similarly you can go for 0 comma 3 number of heads is 0, and the difference in absolute that is going to be 3 that is possible with the tail tail tail. You see that there are totally 8 possible outcomes. So, one we have 3 other we have 3, and other we have one and one so, that total is going to be 8.

Therefore now we can go for finding out the joint probability mass function of x comma y using this box. That is when x takes a value, when x takes a value 0 1 2 or 3, and y takes a value 1 or 3. We can make a table x takes a value 0 y takes a value one that is nothing therefore, the probability is 0, when x takes a value 1, y takes a value one that is a 3 possibilities. It is a unbiased coin therefore, the probabilities going to be 3 by 8.

When x takes a value 2 and y takes a value one there are 3 possibilities, therefore, this is going to be the 3 by 8. When x takes a value 3 y takes a value one and nothing, therefore, no possible outcomes therefore, empty set probability of empty set is 0. Similarly 0 to 0 comma 3, that only one possibility so, 1 by 8, 1 comma 3, there is no possibility therefore, it is 0 and 2 comma 3, there is no possibility therefore, it is 0 and 3 comma 3 is only one possibility, therefore 1 by 8.

If you had all the values 3 by 8 plus 3 by 8 plus 1 by 8 plus 1 by 8, that is going to be 1. If you make a row sum or column sum, you will get the marginal distribution, and if you had those values again you will get the one. So, it is 0 plus 1 by 8 1 by 8, 3 by 8, 3 by 8, 1 by 8, if you add up all these values it is going to be 1. Similarly if you add 3 by 8 plus 3 by 8, that is 6 by 8 1 by 8 plus 0 plus 0 plus 1 by 8 that is 2 by 8.

So, if you add 6 by 8 plus 2 by 8 you are getting 1; that means, the probability mass function of x takes a value small x, that is going to be for different values of x it is 0 1 2 and 3. So, it is going to be for x takes a value 0, that is 1 by 8 for 1 3 by 8 for 2 3 8 and 3 by 8. So, this is going to be a probability mass function of x. Similarly, one can make a probability mass function of y so, different values of y is going to be 1 and 3. So, for one it is 6 by 8 for 3 it is 2 by 8 so, this is a marginal distribution of y.

So, from this page one can get the joint probability mass function of x comma y, from the joint distribution, you can always get the marginal distribution of x and y separately. Or you can find out the marginal distribution from the random variable x itself, you do not need finding the joint distribution then the marginal of x.

You can find the way you have defined x, you can directly get the probability mass function of x. But here what am saying is if you know the joint distribution, you can always get the marginal distribution the other way or the converse is not drawn; that means, from the marginal one cannot get the joint always, whereas, from the joint distribution you can always get the marginal.

Therefore here we get the probability mass function of x and y from the joint probability mass function of x comma y, this easiest example. Since both the random variables are of the discrete type we are able to give the joint probability mass function of the random variable x comma y. So, with this example let me complete the joint probability mass function. In the next class we will go for when both the random variables are of the continuous type, then one can define the joint probability density function that we will do it in the next class.