Probability and Statistics

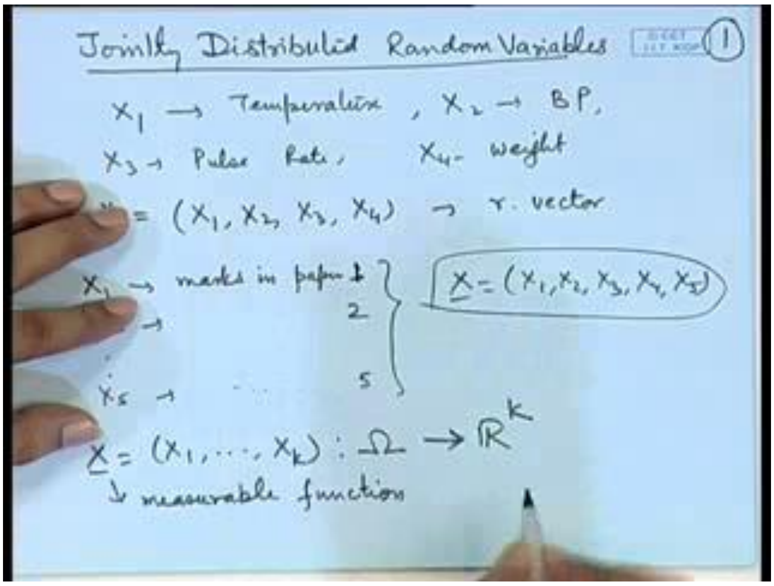
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 35

Joint Distributions – I

So far we have considered the phenomena where if a sample space is given we are considering the function which is considering the mapping sample is place to the real line. So, we are considering one characteristic at a time; for example, it may be heights of the student, it may be marks of students in a text paper, it may be the life electronic equipment. However; so this constitutes the values of a single random variable x.

(Refer Slide Time: 01:21)

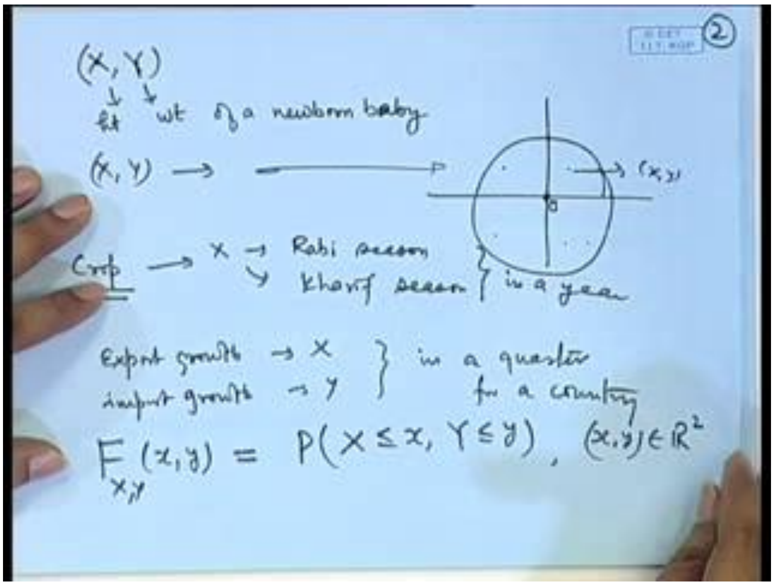


So, we have so for concentrated on the distribution of the random variable x, many times we are not having the luxury of considering only one characteristic, but several characteristics such as a patient goes to a doctor for a medical checkup, the doctor takes his say temperature; he may take his blood pressure, he may take his pulse rate and he may record his say weight. So, for different patients 4 quantities X 1, X 2, X 3, X 4 are recorded. So, here this constitutes say random vector or jointly distributed random variable X 1, X 2, X 3, X 4. So, X is called a random vector. We may be looking at say X 1 as marks in paper 1; X 2 as marks in paper 2; say X 5 marks in paper 5.

So, students are studying 5 subjects in a particular semester, and each of them will get different marks in different test papers. So, for each student if we record X is equal to X 1, X 2, X 3, X 4 and X 5 then this is a random vector or a jointly distributed random variables X 1, X 2, X 3, X 4 and X 5. So, in general if I am considering X as X 1, X 2, X k then this is a function from omega into the k dimensional Euclidean space R k, the random variable X was a measurable function from omega into r that is one dimensional Euclidean space.

So, a higher order random variable or a random vector is a function from it is a measurable function from omega into R k. So, X is measurable function that we have to ensure in order that the probability functions of X are well defined.

(Refer Slide Time: 03:49)



For convenience we will restrict attention to two dimensional case in the beginning; so let us consider jointly distributed random variables X, Y. So, here X could denote the height of a new born and Y could denote the weight of a new born baby.

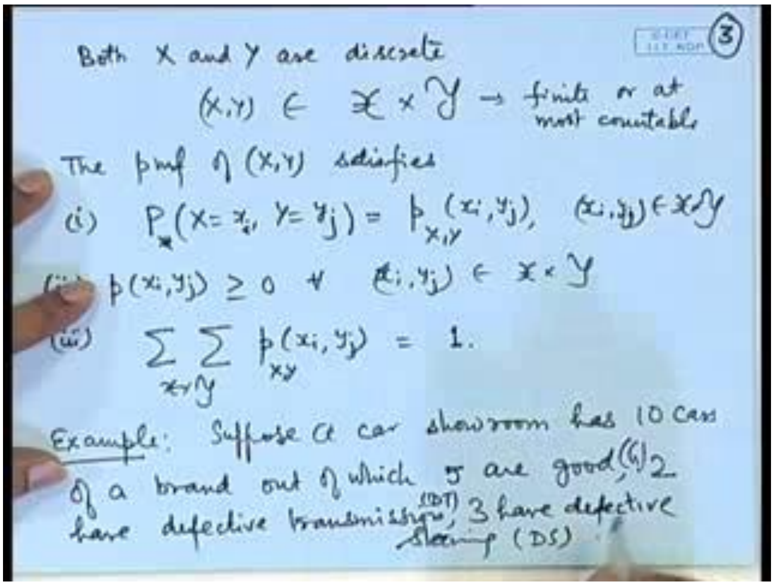
X, Y could denote the coordinates of a dark hitting a target. So, suppose this is the target we consider it as origin and the dot may it anywhere, so the coordinates X Y of the dot hitting may be considered as a jointly distributed random variable. Suppose we are considering crop in a area, so crop in a say Rabi season and in a Kharif season in a year; we may consider say export growth say X and say import growth Y in a quarter of year for a country. So, these are all examples of jointly distributed random variables.

Now, naturally the question arises that how do we evaluate the probability distribution of X Y or how do we define the probability distribution of jointly distributed random variables? So, we have seen in the case of one variable, the description of the probability distribution depends upon whether the random variable is discrete or continuous. So, if the random variable is discrete, we have a probability mass function; if the random variable is continuous, we have a probability density function of course, for any type of random variable we have the facility of using a cumulative distribution function, but that is not useful all the time.

So, when we have a jointly distributed random variable, then there can be several cases; X may be discrete, Y may be continuous both may be discrete both may be continuous, X may be discrete Y may be a mixture. X may be a mixture random variable Y may be continuous etcetera. So, in each case the description of the distribution may be of different nature; however, one may make use of the joint cdf that is probability of X less than or equal to x, Y less than or equal to y, this is defined for all x, y in the two dimensional plane.

So, we will look at the properties of the cdf later on. Firstly, let us consider the case when both x y may be discrete or both x y may be continuous.

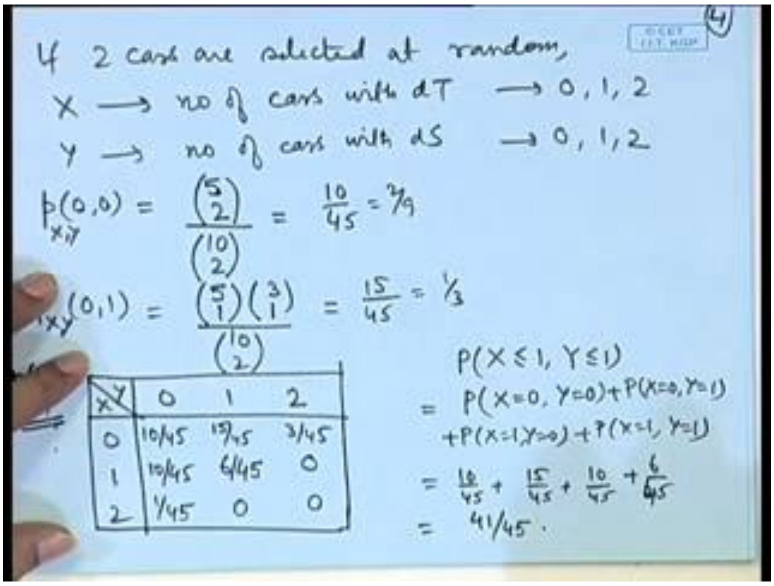
(Refer Slide Time: 07:03)



So, if both X and Y are say discrete; that means, the values that X, Y take in a space x cross y, this is finite or at most countable. So, in this case the probability mass function of X, Y this will satisfy. So, P x probability X equal to x, Y is equal to y, let me put here X equal to x i Y is equal to y j, because the values are at most countable. So, we can enumerate them this is P X ,Y x i, y j for all X y xi, yj then this function is always non negative for all x i, y j and the sum over all the values of the probability mass function is 1.

Let us take one example; here suppose a car show room has 10 cars of a brand out of which 5 are good, 2 are having say defective transmission and 3 have say defective say steering. So, we call it say DT and DS and good are G.

(Refer Slide Time: 09:34)



Now, if 2 cars are selected at random; let X denote the number of cars with defective transmission and Y the number of cars with defective steering mechanism.

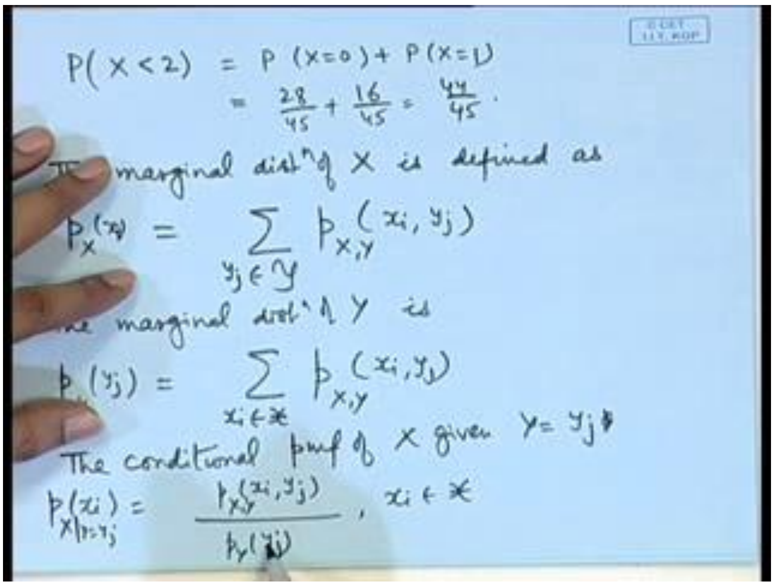
So, here X, Y is a discrete random vector, both X and Y can take values 0, 1, 2. So, the probability distribution of X, Y can be described; what is the probability that X equal to 0, Y is equal to 0? Now this is possible if both of the selections are made from the good 5 cars out of the total selections from 10. So, this is equal to 10 by 45 are 2 by 9. Similarly we can calculate probability X is equal to 0 and Y is equal to 1. So, here it means that out of one good; out of 5 good 1 good car has been selected, and out of 3 cars with defective a steering mechanism 1 has been selected, out of total 2 selections of 10 so that is equal to 15 by 45 or 1 by 3.

In a similar way we can calculate all other probabilities and we can describe the probability distribution in the form of a tabular representation, the probability X equal to 0, Y is equal to 0 is 10 by 45, the probability X equal to 0 Y is equal to 1 is 15 by 45, the probability that X equal to 0 Y is equal to 2 that can be calculated to be 3 by 45, the probability that X equal to 1 Y is equal to 0 can be calculated to be 10 by 45, X is equal to 1 Y is equal to 16 by 45, X is equal to 1 Y is equal to 2 is not possible because total selections are only 2; so this is 0; probability that X equal to 0 Y is equal to X is equal to 2 Y is equal to 0, 1 by 45 again 2 1 and 2; 2 are not possible.

So, this gives the joint distribution P x y of the random variables X Y; you can look at various probabilities related to random variables X Y from here for example, if we ask a question what is a probability that X is say less than or equal to 1, Y is less than or equal to 1 then this is corresponding to probability of X is equal to 0 Y is equal to 0, probability X is equal to 0 Y is equal to 1, probability X equal to 1 Y is equal to 0 and probability X equal to 1 Y is equal to 1.

So, this is equal to 10 by 45, plus 15 by 45, plus 10 by 45, plus 6 by 45, which is equal to 41 by 45.

(Refer Slide Time: 13:41)



We may answer some other questions from here regarding the probabilities, for example if we ask what is a probability X is say less than 2 that is probability X equal to 0 plus probability X equal to 1, now this relates to the probabilities of 1 random variable when the joint distribution is given.

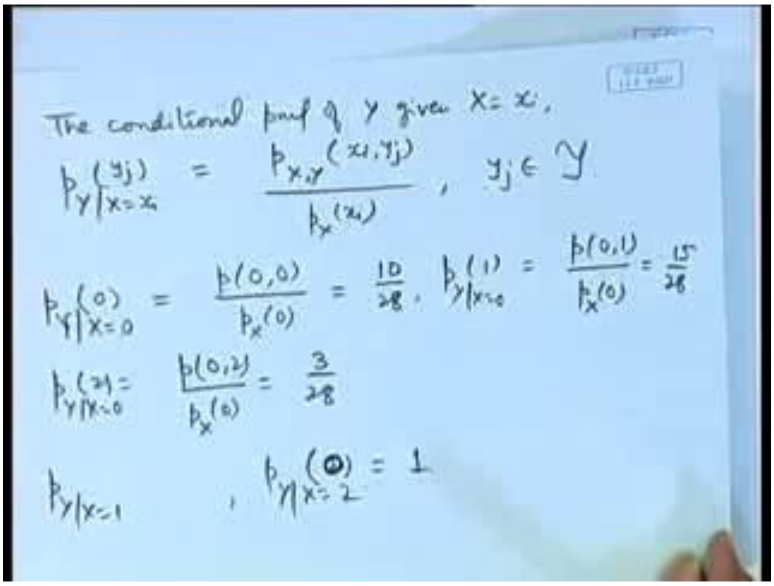
Now, notice here that if I some row wise then this will give probability X equal to 0 Y is equal to 0, probability X equal to 0 Y is equal to 1, probability X equal to 0 Y is equal to 2 that means, this will give the distribution of X that is equal to 28 by 45, 16 by 45 and 1 by 45. Similarly if I add the columns here I will get the distribution of Y. So, that is 21 by 45, 21 by 45 and 3 by 45.

So, now probability X equal to 0 and probability X equal to 1 can be easily obtained as 28 by 45 plus 16 by 45 that is equal to 44 by 45; this process of adding row wise or column wise this gives rise to the individual distributions of X and Y, they are known as marginal distributions. So, in general the marginal distribution of X is defined as this is obtained by adding the values of y over the range of this. Similarly the marginal distribution of Y is obtained by adding the joint probability mass function with respect to x i.

So, we can answer all the probability statements regarding individual distributions of X and Y also from the joint probability mass function. Now there is one more thing when we have 2 random variables, we may also look at the conditioning events for example, if y is equal to 2 is given, if y is equal to 1 is given what happens to the probability distribution of X, this is known as the conditional distribution. So, we may define the conditional probability mass function of X given a value say Y is equal to y j. So, this is given by the joint distribution of X and Y at x i y j divided by the marginal of this where x i varies over x.

To see that it is a valid probability distribution, let us sum over all the values of x i. So, if we sum over all the values of x i, the numerator becomes the marginal distribution of y j which is the denominator. So, this will become 1.

(Refer Slide Time: 17:24)



Similarly the conditional probability mass function of Y given X is equal to x i. So, that is equal to P Y given X equal to x i is add the value y j it is defined to be the joint distribution divided by the marginal distribution of x i for values of y j over y.

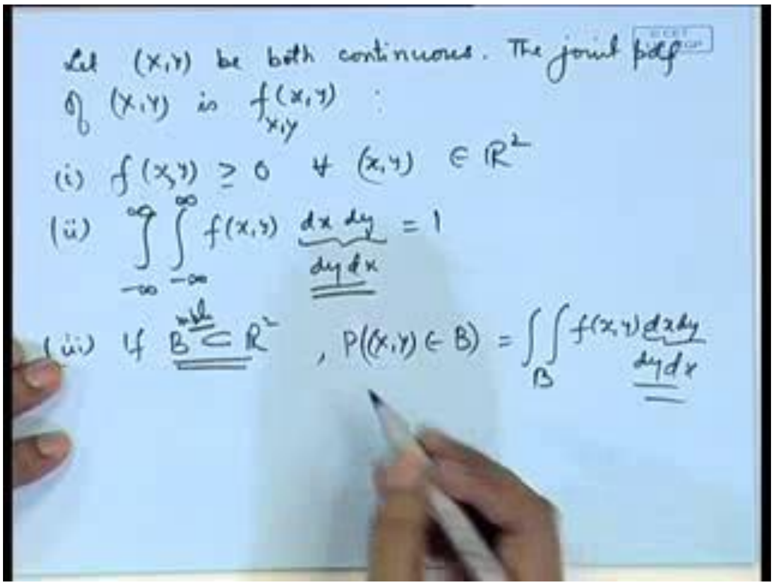
Once again you consider that it is a valid probability distribution, if we sum over the values of y j the numerator here will become P X x i which is same as the denominator so it will give the value 1. So, in this given problem let us look at the conditional distribution say probability of Y given X is equal to 0. So, the values of Y are 0 1 and 2 what is a probability that Y is equal to 0 given X equal to 0? So, it is P 0, 0 divided by P X 0.

Now, P 0 0 is 10 by 45, and P X 0 is 28 by 45. So, this becomes 10 by 28; what is P Y 1 given X equal to 0? This is P 0, 1 divided by P X, 0. So, that is equal to 15 by 28 similarly P Y, X equal to 0 at 2, this is P 0, 2 divided by P X 0 that is 3 by 28. So, you can see here the sum of the 3 probabilities 10 by 28, 15 by 28, and 3 by 28 gives 1. In a similar way one may calculate probability distribution of say Y given X equal to 1, probability distribution of y given X equal to 2 notice here that if I say X is equal to 2 then only Y is equal to 0 is possible.

So, probability of Y is equal to 2, given X equal to 2 that will be 1 this is a degenerate distribution, this is P 2 2 divided by probability equal to 2 that is 1 by 45 divided by 1 by 45, because if X equal to 2 is fixed, then Y cannot take any other value except 0 sorry this is not 2 this is 0, because the total number of cars that we are purchasing is 2 therefore, if one of them is having a value 2 then the other value has to be 0.

Now, if both the random variables are continuous then we have joint probability density function.

(Refer Slide Time: 20:34)



So, let us consider let X and Y be both continuous. So, the joint probability density function of X, Y is f x, y this is satisfying that it has to be a non negative function, the integral over the whole range with respect to both the random variables is 1. So, in place of d x d y one may write d y d x also and if B is a measurable subset of R 2, then probability of X Y belonging to B will be given by the integral of the density over the set B integration may be in any order. If the probability density function is given, he may be able to answer any probability statement related to the distribution of X, Y the joint distribution or marginal distributions of X and Y.

Thank you.