Multivariate Procedures with R

Prof. Shalabh

Department of Mathematics and Statistics

IIT Kanpur

Week – 05

Lecture – 22

Probability Functions for Continuous Bivariate and Multivariate Random Variables

Hello friends, welcome to the course multivariate procedure with R. So, you can recall that in the last lecture we had talked about the concepts of probability density function, probability mass functions, cumulative distribution function etc. Now in this lecture I am going to extend them to a bivariate and multivariate setup. Now if you try to recall that we had discussed that we have two types of random variable discrete and continuous. And when we are talking about their probability function for example, probability density function for continuous random variable and probability mass function for the discrete random variable then they have different types of mathematical properties that we have understood. So, now when we are going forward and when we are trying to discuss the bivariate or say multivariate random variables or random vectors then we will have the concept of probability density function and probability mass function for this bivariate random variables and multivariate random vectors.

So, I can have here both the concept I can discuss here both the details, but as I said in a couple of lectures that in this lecture in this course my objective is to bring you all to a multivariate setup. And in order to bring you to the multivariate setup I am trying to start from the univariate then bivariate and then multivariate so that there is a smooth sailing. And all these topics which I am doing here although they are related to the probability theory and other topics, but they are going to help us in the development of the tools for the multivariate procedures in the forthcoming lectures. So, I have chosen here some selected topics which I am trying to explain you here.

So, now given the forthcoming lectures in the future I am going to use mainly the continuous random variable. The type of tools which we are going to discuss about multivariate analysis like as discriminant analysis, classification analysis etc. I am going to develop them for the continuous random variable. So, that is why in order to save the time I am going to consider only here the case of continuous random variable. And once you understand this continuous case these concept can be directly converted into the discrete case also without much difficulty.

So, this is what you always have to keep in mind when you are trying to attend this course. So, now in this lecture I try to extend the definition of say this probability density function of a univariate random variable to the setup of bivariate random variable. So, let us begin our lecture and try to understand it. So, as we had discussed in the case of univariate continuous random variable that we have the probability density function. Now once again we are going to concentrate only on the continuous random variables but now we have two continuous random variables and we suppose we denote them by X and Y.

So, definitely then now the probability functions in such cases is going to be affected by the behavior of X and Y both and that is why we call it here as a joint probability distribution. And if you try to recall that in the case of univariate we were trying to find out area under a curve like this was to compute different types of probabilities. But now when you have two random variables which are jointly affecting the probability of an event then essentially what is happening that we need to calculate the probability that X and Y assume a value in any given region of two dimensional space. For example now there will be suppose here two random variables one is taking value between A and B and Y is taking value between say C and D. Then now you have to compute the probability of occurrence of say X and Y when X is lying between A and B and Y is lying between C and D.

So, analogous to the definition of probability density function for a single continuous random variable we can also extend and define the joint probability density function over the two dimensional space. So, now if you try to recall in the earlier case we had defined the probability density function by these two condition f(x)>=0 at all point of x and this integral of $f_X(x)\ dx$ should be equal to 1. Now I try to extend these two condition to the setup of a joint probability density function for two continuous random variable X and Y. So, it is now indicated here say f(x,y) and so earlier it was f(x) now it is extended to f(x,y) means the random variables are denoted by upper case alphabet and their values are indicated by lower case alphabets like this. So, now the first condition that f(x) should be greater than or equal to 0 at all points, can be extended to a bivariate setup as $f_{X,Y}(x,y) \geq 0$ for all values of x and y and the condition that $\int_{-\infty}^{\infty} f(x)dx = 1$ can be extended to a bivariate setup by writing $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,y}(x,y)dxdy = 1$.

And earlier we had reason we have taken the region for x to be here R now we have here two random variables both are belonging to here R so this (x,y) will belong to say R x R or say R^2. So, we say here that for any region A belonging to R^2 of a two dimension space if you want to compute the probability that x and y lies in this region A this can be obtained by integrating the $\iint_A f(x,y)\,dxdy$. So, typically this f(x,y) is defined over all of two dimension space by assuming that $f_{X,Y}(x,y)=0$ for all point for which f(x,y) is not specified. So, this is simply an extension of the definition of the univariate probability density function to a bivariate density function. So, now for example if you want to compute any probability in the case of joint probability distribution of two random variables so now if you have a two random variable x and you want to know the probability that x lies between small x1 and small x2 and capital Y lies between small y1 and small y2 then this can be obtained by the integral $\int_{y1}^{y2}\int_{x1}^{x2} f_{X,y}(x,y)dxdy$. So, that is straightforward.

Now in case if I ask you that you have here say p random variables right then if you want to find out here x1 say less than x1 say here y1 or let me write down here like this suppose you have here the random variables here x1,x2…xp and suppose these values are lying between for which you want to compute the probability (a1, b1), (a2, b2),… (ap, bp) and so on then what will be this will be the integration over say here $\int_{a1}^{b1}\int_{a2}^{b2}\dots\int_{ap}^{bp} f(x1,x2,\dots xp)dx1dx2\dots dxp$ like this. So, this is my objective this is what I want to show you right. Now when you are trying to have the concept of joint probability density function then there is an associated concept of marginal probability distribution. What is the difference between the two? Suppose you have two random variables which are going to affect the probability of an event and their behaviour is actually joint that both of them are going to affect the probability of an event. Now you want to know that what is the behaviour of individual random variables.

It is something like if I take a very crude example that suppose in a class a group of students are assigned a project that means the class is divided into different groups and every group has suppose 5 students and now every student is doing a group project. So, when the project outcome comes then it is very difficult to decide that what is the contribution of an individual student but we see only the final outcome. So, the final outcome is something like joint probability that it is the effect of all the 5 students but if I want to find out the effect of only one student then that will be equivalent to marginal distribution or say marginal contribution or equivalently marginal probability distribution. So, if I consider here two random variable X and Y so I can find out here marginal probability distribution from the joint probability distribution function and exactly in the same way as we try to find out the probability function. If it is continuous then integral and if it is discrete then here is sum.

So, an analogous approach is used to determine the marginal probability distribution from the joint probability density function. What exactly you want to do? You want to remove suppose if I want to find out the marginal of here X like this. So, what essentially I have to do? I have to remove the effect of here Y from here. So, now how to get it done? In order to remove the effect mathematically how what you can think of. You see whenever we want to do something we have to think of a mathematical procedure that can execute it.

So, in case if I try to integrate out the effect of Y, Y's effect then the remaining part will be treated as an effect of only X. So, that is what we try to do in the case of marginal probability distribution. So, if there are more than one random variables that are associated with the random experiment we distinguish between the joint probability distribution of X and Y and the probability distribution of each variable individually, that mean the probability distribution of X, probability distribution of Y and we have a joint probability distribution of X and Y. And this individual probability distribution of a random variable they are referred to as marginal probability distribution. For example, marginal probability distribution of X, maximal probability distribution of probability distribution of Y.

So, in general the marginal probability distribution of X can be determined from the joint probability distribution of X and other random variables in general. So, let me try to show you here in the case of bivariate which can be very easily extended to a multivariate case. So, if I have here bivariate probability density function $f_{X,Y}(x, y)$. Now you want to find out here marginal distribution of X which is indicated here by f_X. So, now I am trying to integrate out the effect of Y just by integrating the joint probability density function over the range of Y and so it will be the range of here Y and here d Y right.

So, now I believe that the effect of Y has been taken out and now whatever is the effect remaining that is that belongs to X. And similarly if you want to find out the marginal distribution of Y then it is something like you try to take here the joint frequency function and try to integrate it out over the range of here X like this integral over $\int f_{X,Y}(x, y)dx$. And now if I ask you suppose if you want to extend it to a multivariate setup suppose if you have here three variables x, y, z and based on that you have here a probability density function is a joint density function $f_{X,Y,Z}$. And suppose if you want to find out here the effect of here x f_X. So, what it now will become? So, $f_X(x)$ will become here integral over Y integral over $\int f_{X,Y,Z}(x, y, z)dydz$ now this will be over dy and dz.

And similarly if I try to take here more than three variables suppose if I have random variable x1, x2, x3, x4, x5 and suppose I want to find out here the joint effect of x1, x2 by removing the effect of x3, x4, x5. So, for that obviously you will have here a probability density function f(x1, x2, x3, x4, x5) like this. And now you want to obtain here suppose f(x1, x2). So, what are you going to do here? You try to write down this function and try to integrate it over say $\int f(x1, x2, x3, x4, x5)$ dx3 dx4 dx5 . So, this is how you can extend this univariate concept to a multivariate concept.

And similarly we also have here concept of conditional density function or conditional distribution. So, when we have a joint density function say f of x y then if I want to find out the individual contribution of x and y then we use the concept of marginal distribution. But suppose if I want to see the effect of one variable when the effect of another variable is fixed then how to find out this effect? So, this is obtained by the concept of conditional probability distribution. So, if I have here two random variables x and y then the conditional distribution of x given capital Y equal to y is obtained here like this. So, what is happening? The value of y here is fixed.

Now I want to see the behaviour of x that how the probabilities are changing. So, this is indicated by here the random variable like as here x given y is equal to y or it is written as a capital X this vertical line and then here is small y. So, now you want to find out the conditional density of x. So, it is written here the random variable here is x given y and then here this is here x. So, this will be here joint probability density function divided by the marginal density function.

So, here in the numerator this is joint probability density of x and y and in the denominator, this is the marginal density of y. So, this is how you can obtain the distribution of the conditional probabilities. And similarly, if you want to change the role of x and y then the conditional distribution of y given x is here like this. You can see here now here the random variable is y given x and the values which are given is here y and it is here the joint PDF in the numerator and marginal density of x in the denominator right. This concept can be extended to a multivariate setup also right.

Suppose if I say there are here suppose here 4 random variables x1, x2, x3, x4 and suppose I want to see the joint effect of x1, x2 condition that x3 and x4 they are here fixed. So, similar type of definition can be given here also in the numerator this will be x1, x2, x3, x4 and in the it is going to be here $f_{X3,X4}(x3, x4)$ like this right. So, now you can see here this concepts can be extended to a multivariate setup without any problem. Now let me try to take some examples to explain you how these probabilities can be computed. They are very simple example just to give you an illustration.

Suppose we have a joint probability density function of x and y which is given here by this function which takes really twice of exponential of minus x into exponential of minus 2 y when x lies between 0 and infinity and y lies between 0 and infinity and 0 otherwise elsewhere. Suppose if you want to compute the probability that x greater than 1 and y less than 1. So, this type of integral with the range of x between 0 and 1 and that means range for y as 1 to infinity this joint PDF this can be solved. I am not going to solve here this integral, but anyway you can obtain this value. Similarly, if you want to obtain here the probability of x less than y.

Now it is a joint variation. So, you can write down here the limits of x as 0 to infinity and probability of and the limits for y as 0 to y write down the PDF and solve it. This will come out to be here 1 upon 3. And similarly, if you want to just find out that the probability of x is less than c that mean probability and that mean the y is lying between 0 to infinity you can have this integral range for x and 0 to infinity range for y the joint PDF and then you can solve this integral without any problem. So, now you can see here different types of this probabilities can be computed from this joint probability density function. Now we consider another example to show you the application of marginal distribution.

Suppose our joint PDF is given by here like this that f(x , y) =2 if 0 < x < y < 1 and f(x , y) =0 otherwise. If you want to find out here the marginal of here y then it is here you have to integrate over the range of x which is here 0 to y and then this joint PDF and if you solve this it will come out to be here twice of y and similarly if you want to find out here the marginal of here x then you have to integrate over the range of here y over the PDF and this will come here like. So, you can see here it is only an integration problem that you have learnt in your undergraduate. And similarly, if you want to find out the conditional density function in such a case. So, now you already have found here these marginal densities.

So, by the definition of conditional density function of x given y this is here joint PDF divided by marginal PDF of y. So, it will come out to be 2/2y which is here like this $\begin{cases} \frac{1}{y}, 0 < x < y \\ 0, otherwise \end{cases}$. And if you similarly if you want to find out the conditional density of y given x this is the joint PDF of x and y divided by the marginal of here x which is equal to here like this you already have obtained this marginal PDF and the joint PDF here given by 2 and this is the marginal PDF. So, it will be here like this. So, you can see here finding out this conditional and marginals are equivalent to finding out any integral problem that is all.

Now, after this we have another concept of stochastic independence which is in usual word we call it as independent random variables or independence. So, what is the basic concept of independence? That if you have more than one random variable suppose I have here two random variables x and y and suppose if the occurrence of x does not affect the probability of y and if the occurrence of y does not affect the means occurrence of x or the probability of x that means x and y they are not affecting to each other then we say that they are independent. So, two continuous random variable x and y are said to be stochastically independent stochastic means random. If the joint density function of x and y can be expressed as the product of their marginal density function. So, the joint density function of x and y can be expressed as the marginal density function of x and marginal density function of y.

And the same concept can be extended to discrete random variable also that the joint probability that $X = x_i$ and $Y = y_j$ this can be expressed as the product of their individual probabilities that probability $P(X = x_i)$ into probability that $P\left(Y = y_j\right)$ for all values of i,j. And the same condition for independence in terms of a cumulative distribution function can also be expressed. In terms of joint CDF for example, here like this F(x, y) the random variable x and y are independent if the joint CDF can be expressed as the product of their product of the CDF of individual variable that is F(x, y) can be expressed as the product of F(x) and F(y). So, if you want to see through an example I can consider here joint probability distribution function here like this and I try to find out here the marginal density of here x which is here integral over the range of y and then the same integral. Now you can solve this integral it will come out to be here like this.

Similarly, if you try to find out the marginal density of y it will be integrated over the range of here x into with this probability density function and if you try to solve this integral this will come out to be here like this. So, now you can see here that the product of this and this, this is the same as this given joint density function that is $f(x, y) = f(x)\, f(y)$ and this implies that x and y are independent. So, this is how we try to find out that if any set of random variable they are independent or not. Now the same concept can be extended to a multivariate setup also. If I see here if I have got here p random variables then if they can be expressed as $F(x1)\ F(x2)\ \ldots\ F(xp)$ and the same concept is true for the CDF also up to here F(xp).

So, now you can see here how I have extended the univariate and bivariate concept to a multivariate setup, but definitely I will try to introduce you these concept in a multivariate setup when we try to consider the multivariate normal distribution. But here I would like to conclude it by defining the way we are going to consider the random variables or a random vector. So, suppose we have p random variables which are indicated by x1, x2,… xp and now they can be combined in a vector which is a matrix type vector. It is not a data vector because what I am trying to say. So, all these variables can be combined in a random vector of order say p cross 1 like this $\underline{X}$ and these are here this random variables X1, X2,… Xp and they can be also expressed here as say x1, x2,… xp transpose and when you are trying to find out the space of it this is going to be the set of n tuples.

And now you can define here pdf something like here (X1, X2,… Xp) = (x1, x2,… xp) or in brief it can be written here as say f(x) and like here at as this and then based on this you can define different types of marginal density function, conditional density function etcetera right. So, in general we will try to denote the pdf of the random vector x by here like this and similarly we can define the marginal distribution independence of the random vectors can be decided, but anyway we will try to handle it when we try to consider a multivariate normal density function right. So, now we come to an end to this lecture. So, now you can see what we have done in this lecture. Earlier lecture we introduced the univariate random variable.

Now in this lecture I extended the same concept to a bivariate random variable and very silently and gradually I extended the same concept to a multivariate setup. So, we had univariate random variable, bivariate random variable and now we have a random vector. The concept of marginal distributions, the concept of conditional distribution they occur only when we have at least two random variables because you need something to marginalize or something to be conditioned. But the concept of marginal distribution and conditional distribution that I have explained in the context of bivariate random variable they can be extended to the multivariate setup also. In the bivariate setup I had only one variable, one variable for the conditional, but in the multivariate setup I can have two random sub vectors.

When we are trying to talk about the independence we can talk about the independence of the individual components, we can talk about the independence of the random sub vectors also. For example, if I have five random variable x1, x2, x3, x4, x5. Suppose x1 and x2 and another set of variable x3, x4, x5 they are independent. So, all these concepts they are derived from this very simple concept, but they are extended to a situation and they are modified according to the need of the time. So, my request is that you try to pick up any book and try to at least choose some example on the bivariate probability density function and try to just see how you can define them for discrete random variable, how you can define them for the continuous random variable and how you can obtain different types of marginal density function, conditional density function and definitely this will make you more confident.

Once you are confident then nobody can stop you in excelling in the life. So, you try to practice it and I will see you in the next lecture with more topics on this probability function till then goodbye. Thank you.