PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 08

Lecture - 40

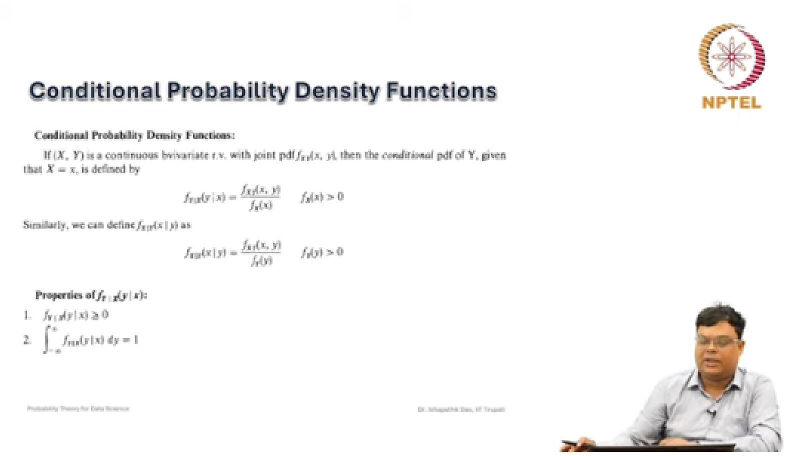
Conditional Probability Density Function

Let us discuss now the conditional probability density function. Similar to the discrete bivariate random variable case, here we will discuss it for continuous bivariate random variables, where both Y and X are continuous. We will look at how the conditional probability density function of Y, given that X = xi, can be found. Here, the definitions are provided for the conditional probability density function. If (X, Y) is a continuous bivariate random variable with a joint probability density function fxy(x, y), then the conditional probability density function of Y given X = x is defined by this relation:

f(y | x) = fxy(x, y) / fx(x), where fx(x) ≠ 0.

Similarly, the conditional probability density function of X given Y is defined by:

f(x | y) = fxy(x, y) / fy(y), where fy(y) ≠ 0.

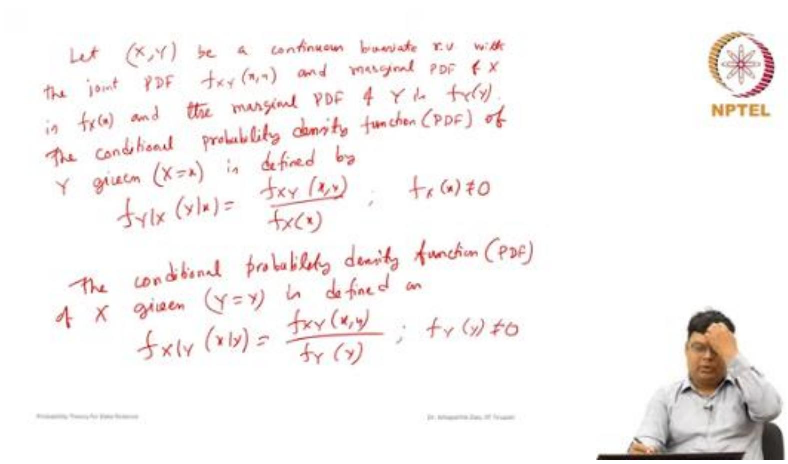


Also, the properties are very similar to the conditional probability mass function. Let us write it down. Let (X, Y) be a continuous bivariate random variable with the joint probability density function. Here, we will talk about the joint probability density function because the variables are continuous, rather than using a probability mass function. We will discuss the joint probability density function fxy(x, y), the marginal probability density function of X, fx(x), and the marginal probability density function of Y, fy(y). The conditional probability density function (pdf) of Y, given X = some point x, is defined by f(y | x). Similarly, when we define the probability mass function, we will use the joint probability density function divided by the marginal probability density function of X.

So, provided that fx(x) ≠ 0, the conditional probability density function (PDF) of X given Y is defined as follows:

f(x | y) = fxy(x, y) / fy(y), provided fy(y) ≠ 0.

So, this is the definition of the conditional probability density function. Some of the properties are very similar to the discrete cases. Since it is a probability density function, for any value of Y, one property is that the total probability must sum to 1. So, for Y given X, let us write down one of the density functions. Y given X should be greater than or equal to 0.

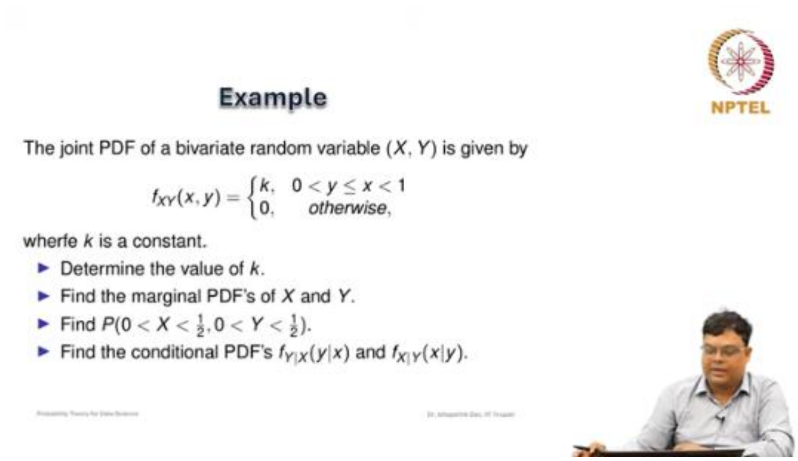


Also, if you fix X for any value of X, the integral of Y given X from -∞ to +∞ should be equal to 1. Because it is a density function, this should be equal to 1, where x belongs to the range of values for which fx(x) ≠ 0. So, x has to be within Rx for this to be defined, and this integration will be equal to 1:

∫ (from -∞ to +∞) f(y | x) dy = 1, for x ∈ Rx, where fx(x) ≠ 0.

Similarly, for other cases, it is defined in the same way. It is very similar to the discrete case, but this is for continuous cases. The definitions apply to continuous cases.

Let's work through a numerical example using the probability density function to find the result for a bivariate continuous random variable. In this example, the joint probability density function of the bivariate random variable (X, Y) is given by this. Let us write down the joint probability density function of (X, Y). For the continuous random variables (X, Y), it is given by fxy(x, y).



This is equal to some constant k, where 0 ≤ y ≤ x ≤ 1. It is equal to 0 otherwise. So, it is non-zero in the domain where 0 ≤ y ≤ x ≤ 1. If you draw the graph, suppose this is (0, 0), and this is (1, 1). This is (0, 1), and this is (1, 0), and this is (1, 1).

Now, if you draw this line, it represents the line where y = x. So, y ≤ x in this region, because here, y = x. This region is where y ≤ x. It is a uniform distribution in this area. So, what is the area of this region?

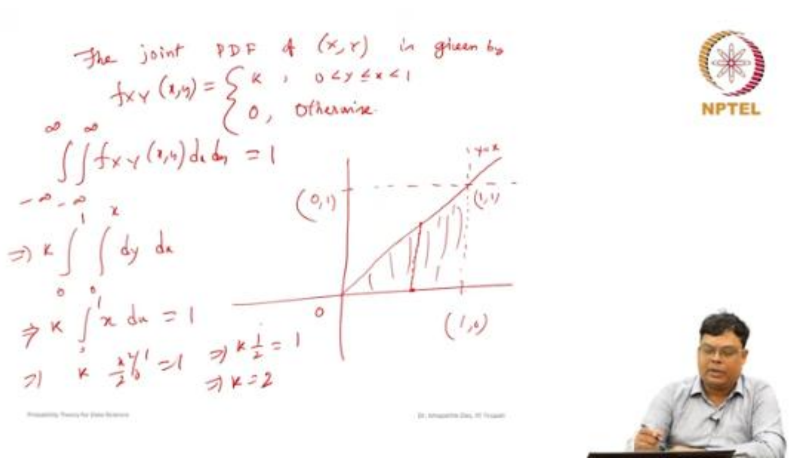
The area of this region should be 1/2 of this square. The area of the square is 1 × 1. Therefore, the area of this region is 1/2. The triangle's area is also 1/2. Now, we need to determine the value of k, which is the constant for the probability density function.

This must be greater than or equal to 0, as k cannot be negative. It has to be greater than or equal to 0. Now, how do we find the value of k? By the properties of the density function, the integral from -∞ to +∞ of the joint probability density function fxy(x, y) with respect to x and y should be equal to 1. This function is non-zero only when y ≤ x and x ≤ 1. So, y can go from a minimum of 0 to a maximum value of x.

This implies that if you integrate with respect to y first, the limits for y can be written as 0 to x, or you can write it the other way as well. Then, this is the constant k, and the integration is done with respect to y first, while x can range from 0 to 1. I want to integrate with respect to y first. For any value of y, if you fix a particular value of x, the limit of y will be determined. Here, it is given that y is between 0 and x. So, the minimum value of y for a particular value of x will be from 0 to x.

The limit of y is from 0 to x, and the limit of x can range from 0 to 1. This integration will be equal to k because I changed the integral to make it simpler; otherwise, the value will remain the same. This becomes an integration of x dx from 0 to 1, with the limit of y being from 0 to x. This is equal to 1, which implies $k * (x^2 / 2)$ from 0 to 1 equals 1. This simplifies to $x^2 / 2$.

Taking the limit, we get 1/2, so $k * 1/2 = 1$, which implies $k = 2$. Therefore, the value of k is 2, and the area is 1/2. Finally, the joint probability density function of (X, Y), denoted as f(x, y), is given by 2 whenever $0 \leq y \leq x \leq 1$. Since it is a continuous random variable, we are writing the joint probability density function as 0 otherwise.

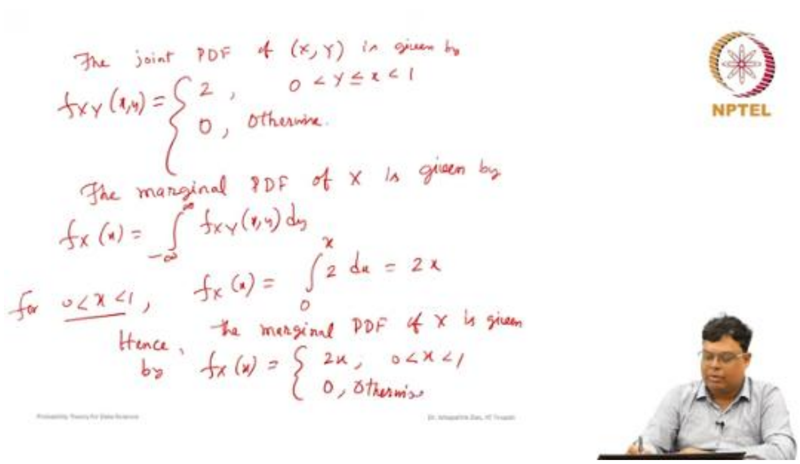


We have found that $k = 2$. Now, we need to find the marginal probability density functions of X and Y. The marginal probability density function of X is given by the definition: $fx(x) = \int fxy(x, y)\, dy$.

But fy(y), sorry, it is the joint density fxy(x, y) dy. We need to integrate the joint density function with respect to the other variable to find the marginal density of X.

Now, this joint density function is non-zero in this region. For x between 0 and 1, any value of y can range from 0 to x. So, the limits are as given. So, it is straightforward. This is nothing but the integral from 0 to x, for $0 < x < 1$. The marginal density function of X, $f_X(x)$, is equal to the integral from 0 to x of 2x dx.

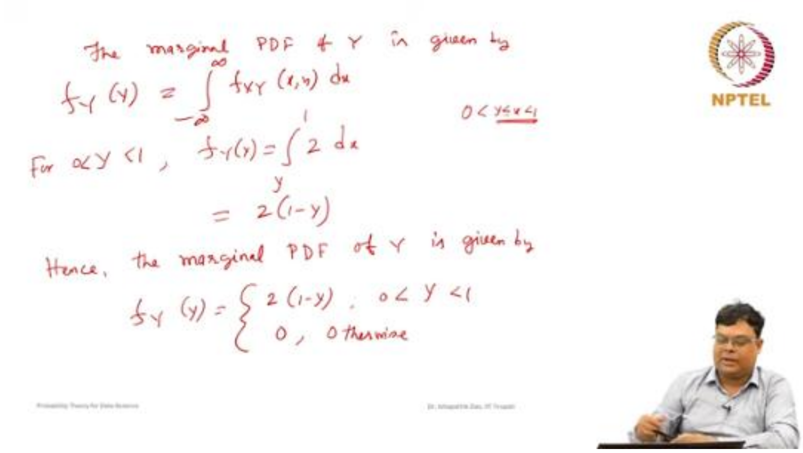
Hence, the marginal probability density function of X is equal to 2x. Whenever $x > 0$ and $x < 1$, for any other value, this density is 0. Therefore, the integration will also be 0, so it will be 0 otherwise. Now, we want to find the marginal probability density function of Y. The marginal probability density function of Y is given by fy(y).



By definition, it is the integral from minus infinity to plus infinity of fxy, with respect to x. So, we have to integrate only with respect to x because we want to find the marginal probability density function of Y. For Y between 0 and 1, as given, the value of fy(y) is non-zero whenever $0 \leq y \leq x$, where x is between 0 and 1. We need to find the limits for this integration. So, the limit of x will be from y to 1 for a given y.

We need to find the limit of x. This is equal to the integral from y to 1, with f(x, y) being constant and equal to 2. So, the result will be $2 * (1 - y)$. Otherwise, this value will be 0. Hence, the marginal probability density function of Y is given by:

Hence, the marginal probability density function of Y is equal to $2 * (1 - y)$ whenever $0 \leq y \leq 1$. For any value outside this region, the function is equal to 0. So, this is the marginal probability mass function. Now, the question is to find the probability of x being between 0 and 1/2, and y being between 0 and 1/2. Let us calculate the probability.



The probability that x is between 0 and 1/2, and y is between 0 and 1/2, is given by the definition: integrate from 0 to 1/2, and from 0 to 1/2, the joint density function fxy(x, y) with respect to x and y. But this function is not non-zero over the entire region. Let us check that. If you draw the graph, suppose this is 0, 1, and this is 1. The points (0, 1), (1, 0), and (1, 1) define the region.

The density function is non-zero only inside this region. Now, if you take this half, the half is here, with x ranging from 1 to 1/2, and y here, from 0 to 1/2. Then, we have to perform the integration over this region. This area will be proportional to 1/2 * 1/2, which equals 1/4, or 1/8, actually. That is why we need to find this.

So, now this will be nothing but the density function, which will be 2 dx dy. For simplicity, let us consider that this density function will be valid for any value between 0 and 1/2. This density function is non-zero in the region where $0 < y \leq x$, and $x \leq 1$. This is given, so in this region, the value is equal to 2. So, let us first take the limit of y.

This will make it a little simpler. For computation purposes, we will use dy dx. This is equal to 2, where x can range from 0 to 1/2. Between 0 and 1/2, for particular values of x, y can range from 0 to x. So, this is from 0 to x, where y can range from 0 to x.

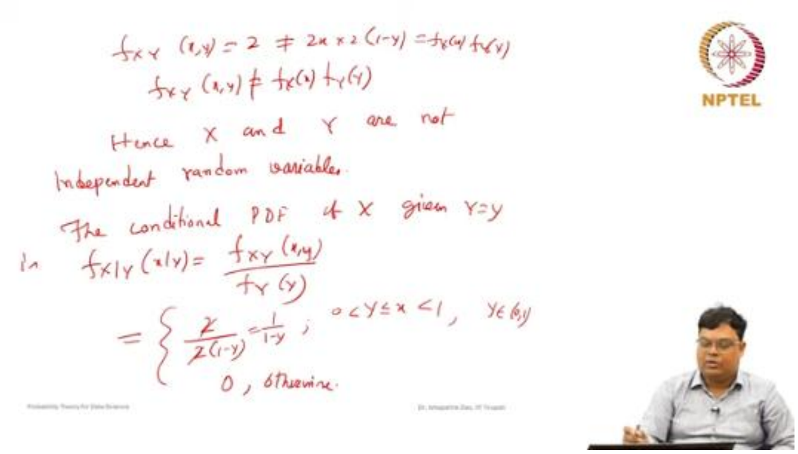
This is equal to 2 * y * x, then dx, from 0 to 1/2. This simplifies to 2 * x^2 / 2, from 0 to 1/2. The 2 cancels out, leaving $(1/2)^2$, which is equal to 1/4. So, the probability is 1/4. So, in this way, whenever the joint probability density function is given, we can find the probability. We have to properly identify the region where this density function is non-zero. Accordingly, we perform the integration to find the probability.

Now, the last question is to find the conditional probability density functions of Y given X and X given Y. By definition, we need to find them. So, note that this random variable is not independent. You can see that the joint density function is constant, which indicates a uniform distribution. Now, for fx(x), it is 2x, and for fy(y), it is 2 * (1 - y). So, fy(y) is 2 * (1 - y). Clearly, fxy(x, y) is 2, which is not equal to 2x * 2 * (1 - y), in general. This is not equal to fx(x) * fy(y).

So, basically, we find that fxy(x, y) is not equal to fx(x) * fy(y) for all x and y. It is not true for any particular values of x and y either. Hence, X and Y are not independent random variables. Independent random variables cannot be assumed here. That is why we cannot simply say that the conditional probability density function will be equal to the marginals.

We must find the conditional probability density function. The conditional probability density function of X given Y is equal to fxy(x, y), assuming that fy(y) is non-zero. This is the definition. Now, note that this will not be non-zero for the entire region. So, this is equal to f(x, y), which will be 2 whenever $y \leq x$ and x is between 0 and 1.

Fy(y) will be non-zero whenever y belongs to the interval (0, 1), as given. For particular values of y between 0 and 1, fy(y) is given by 2 * (1 - y). This simplifies to 2 * (1 - y), and the terms cancel. This is 1 / (1 - y), and it is equal to 0 otherwise. So, just again you can write for simplifying.

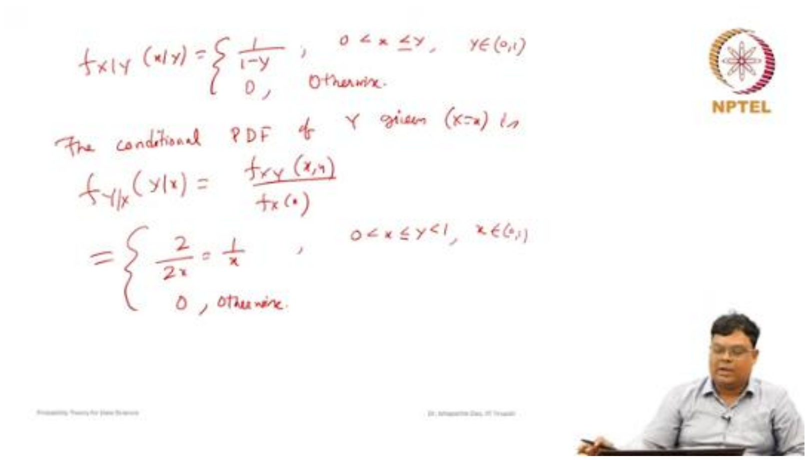


So, the conditional probability density function of X given Y is 1 / (2 * 2), which equals 1 / (1 - y), whenever this is in the region where $0 \leq x$ and $x > y$, and $y < 1$. So, now it is given that we need to find the limit of x only. Since it is constant for x between 0 and y, this is 1 - y, which represents a uniform distribution. For particular values of y in the interval (0, 1), it should not be exactly equal to 1. That is why the open interval is taken.

So, this is 1 / (1 - y) whenever $0 \leq x \leq y$, and it is 0 otherwise. Similarly, the conditional probability density function of Y given X is equal to fy(y), and fy(y) given x. By definition, this is fxy(x, y) / fx(x). This is non-zero when f(x, y) = 2, whenever $0 \leq x \leq y$, and $y < 1$. Apologies for the confusion, but based on the information provided in the problem, this is an open interval where $x > 0$ and $x < y$, and $y \leq 1$.

For x, it lies between 0 and 1. So, this is the density function of Y given a particular value of x. Now, what is fx(x)? fx(x) is the marginal probability density function of X. Here, fx(x) is given as 2x.

Therefore, this is equal to 2x, and it simplifies to 1 / x. It is 0 otherwise. So, now we can write down the conditional probability density function properly. This is the conditional probability density function of Y given X. Hence, the conditional probability density function (PDF) of Y given X, denoted as Y given X, is as follows.



This is equal to 1 / x whenever y is between x and 1. So, y is between x and 1. This is 1 / x. I did it. Oh, this limit is actually... Yeah, so that's why $0 < y \leq x$, and $x < 1$. So, I made this mistake. Sorry, please correct this. Hopefully, you have observed these things.

It is actually just the interchange of x and y. So, $0 < y \leq x$ here. And here, this is y, and this is x. So, finally, we will get that wherever it is $0 < y \leq x$, and $x < 1$, there is a mistake. It is actually just interchanged. So, here, we also have to give this limit: $0 < y \leq x$, where x is between 0 and 1. Any value it can take, then it is 1 / x. Similarly, here, this is 1 / 2, 1 / 2, with x between certain values. Actually, it should be 1 / (1 - y) if 0, with x between y.

So, we need to write it like this: x is between y and 1. Essentially, the important part is the limit of x. x is between y and 1. That is why x will be 1 - y whenever y takes any value between 0 and 1. It will be an open interval from 0 to 1. The limit of x will be from 1 - y to 1. Therefore, it is 1 - y, which represents a uniform distribution in the range y to 1.

Finally, the limit will be written as 1 / x whenever the limit of y is from 0 to x, with x belonging to the range 0 to 1. This will be 0 otherwise. This is the conditional probability density function of Y given X, and this is the final answer. These are the questions provided in this problem. Using the conditional probability density function we discussed, along with the joint probability density function, we can compute probabilities, including their marginal probability density functions.

We also checked whether the variables are independent. In this case, X and Y are not independent random variables. Additionally, we found that although the joint probability density function of (X, Y) is a uniform distribution in this region, the marginals are not uniform distributions. However, the conditional probability density functions are again uniform distributions in a particular region, given X and Y, with those distributions being uniform accordingly.

Hopefully, you have understood this example. Next, we will discuss other topics, such as how we can measure the association between two random variables. For both bivariate and univariate random variables, we have defined some of the moments—first-order moments, second-order moments, and nth-order moments. We will now discuss how we can define these moments.