PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

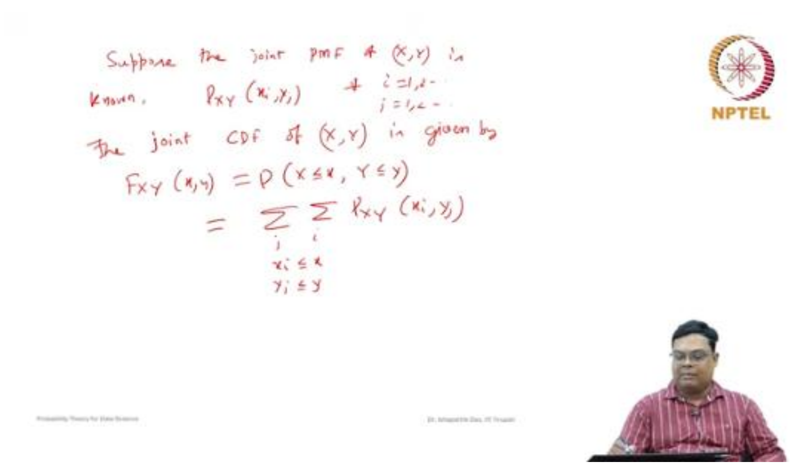
Week - 07

Lecture - 35

Joint Probability Mass Function, Marginal Probability Mass Function, Examples

So, if you know the joint probability mass function, suppose the joint probability mass function of (X, Y) is known. This means Pxy(xi, yj) is given for all i = 1, 2, ... and j = 1, 2, ... The joint cumulative distribution function (CDF) of (X, Y) is given by Fxy(x, y). This is equal to the probability that $X \leq x$ and $Y \leq y$. So, how can we find that?

So, it is actually the summation of all i and all j, Pxy(xi, yj), such that $xi \leq x$ and $yj \leq y$. x and y are fixed, and for all i and j satisfying this relationship, we take the sum to get the cumulative distribution function from the joint probability mass function. It is very similar to the univariate case we discussed. These are some of the properties of the joint probability mass function. Let us discuss one numerical example.



Before that, we need to compute these values, so it is important to clarify this concept and understand how to find them. It is very similar to the univariate case we discussed. These are some of the properties of the joint probability mass function. Let us discuss one numerical example. Before that, we need to compute these values, so it is important to clarify this concept and understand how to find them.

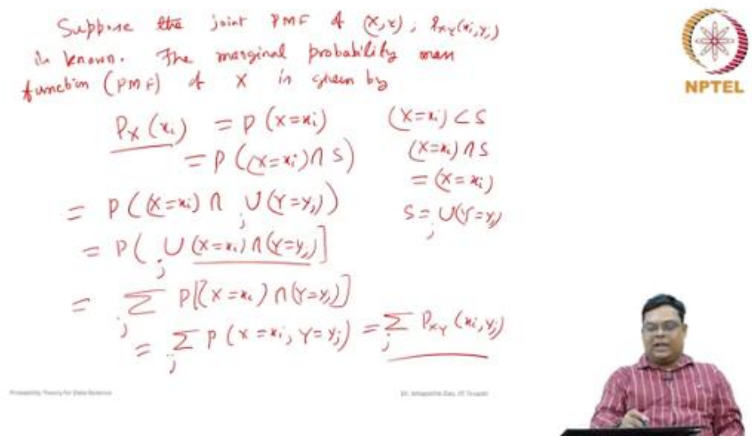
To find the marginal probability mass functions, we consider the subset of the joint probability mass function. This is referred to as a random vector. Since there are only two variables, X and Y, we want to find the probability mass function of X and the probability mass function of Y. So, how can we find the probability mass function of X? You probably know that the marginal probability mass function (PMF) of X is given by F_X(xi), which represents the probability that X = xi.

How do we find that? This can be represented as follows: the event X = xi can be expressed as the intersection of S, since X = xi is a subset of S. Therefore, the intersection of X = xi with the sample space represents the event X = xi. Now, S can be represented as X = xi, which can be expressed as the union of all j, where Y = yj. If you consider Y as a random variable taking values such as y1, y2, y3, and so on, then the union of all possible values that Y can take will be equivalent to S.

S can be represented as Y = yj. Now, using the distributive property, this is equivalent to the union of all j, where X = xi and Y = yj. Then, we use Axiom 3. This is the union of all j, where X = xi and Y = yj. The union is valid because these are all disjoint; y1, y2, and y3 are different, which is why they are disjoint.

This is equivalent to the summation over j of Pxy(xi, yj). This is essentially the summation over j of the joint probability that X = xi and Y = yj. In other words, this is the joint probability mass function of (xi, yj). So, what have we found? To find the marginal probability mass function of xi using the joint probability mass function, you fix a particular value of xi and then sum over the other values of yj. This is the summation over j of Pxy(xi, yj).

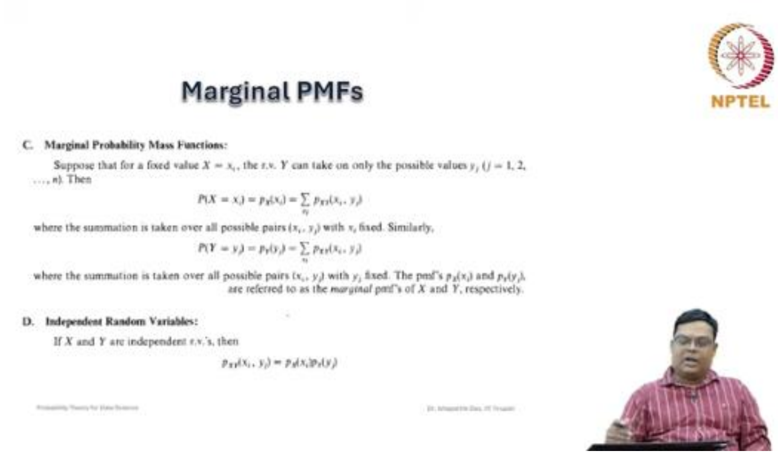
Similarly, the marginal probability mass function of Y, denoted by Py(yj), can be found by summing over all the other variables (xi, yj). This will give you the marginal probability mass function of Y. It is similar to what we did earlier. You fix xi and then sum over the other values of yj. This will give you the marginal probability mass function of Y, denoted by Py(yj).



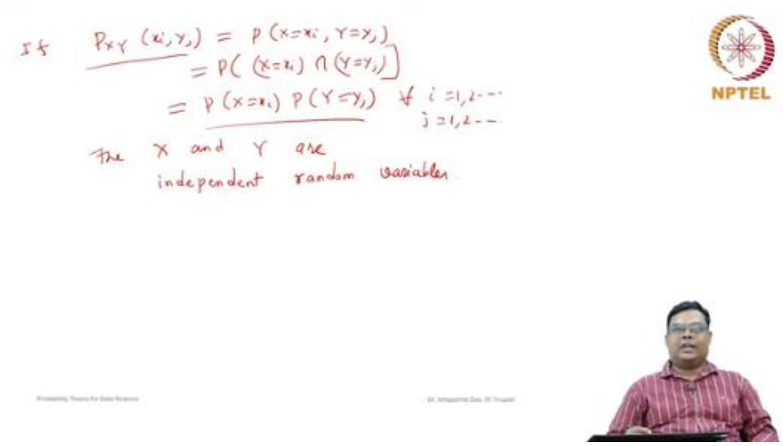
This is the probability that Y = yj. So, that's the concept here. You can see how to find the marginal probability mass function. This is the probability that Y equals the value yj. That is the concept here.

As you can see, it explains how we can find the marginal probability mass function. Suppose that for a fixed value of X = xi, the random variable Y can only take the possible values yj, where j ranges from 1 to n. The probability that X = xi, which is the marginal probability mass function, can be found by summing over all the other values of yj. Similarly, if you take the summation over all possible pairs (xi, yj), with xi fixed, you can find the marginal probability mass function of Y by fixing yj and summing over the other values of xi for Pxy(xi, yj). The summation is taken over all possible pairs (xi, yj) with yj fixed.

The probability mass functions Px(xi) and Py(yj) are referred to as the marginal probability mass functions of X and Y, respectively. The probability mass function Px of xi and Py of yj are referred to as the marginal probability mass function that we mentioned here X and Y respectively. Now, with respect to the marginal probability mass function, we can also represent how two random variables are independent. Pxy(xi, yj) represents the probability that X = xi and Y = yj. This is essentially the event X = xi ∩ Y = yj. Pxy(xi, yj) represents the probability that X = xi and Y = yj. This is essentially the event X = xi ∩ Y = yj.

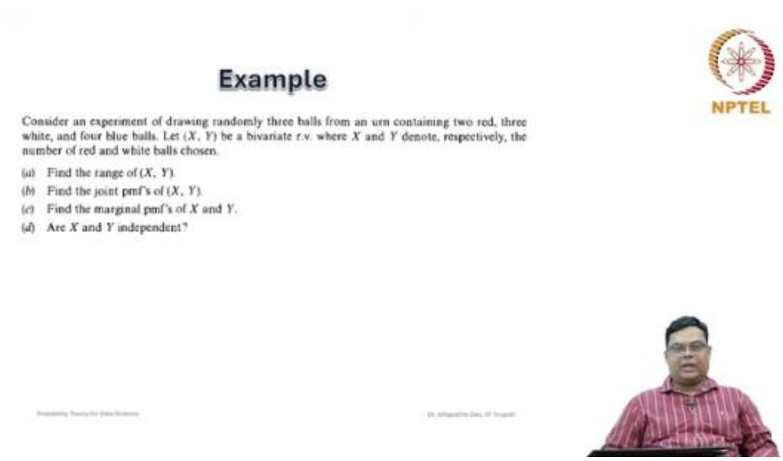


If we find that this is equal to X = xi, then, by the definition of independence, the probability of the intersection of A and B is equal to the product of the probabilities of A and B. If this condition is satisfied, it may not always be the case, but if the condition holds for all i and j, where i = 1, 2, ... and j = 1, 2, ..., then X and Y are independent random variables. If this is satisfied for all xi, it may be true for particular values of xi and yj, but we cannot say that (X, Y) are independent random variables. For independence, we require that this condition is satisfied for any i and j. Only then can we say that X and Y are independent random variables.



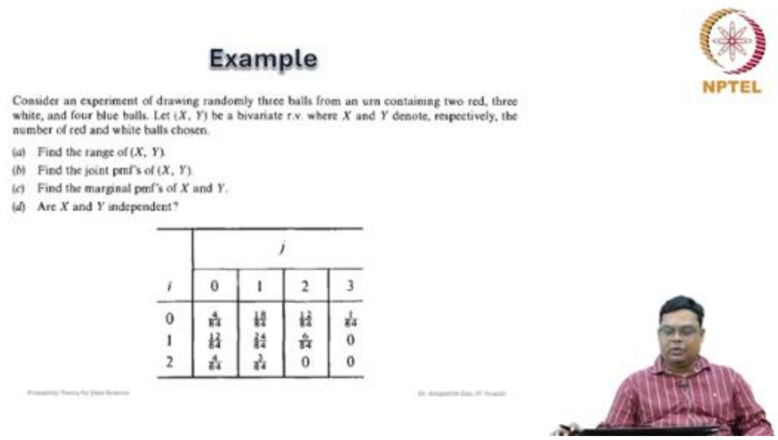
This is the joint probability mass function we discussed, and it shows how we can find the marginals using the joint probability mass function, as well as how we define independence. So, if X and Y are independent random variables, this will be true for all xi and yj belonging to the range of (x, y). We can say this for all i = 1, 2, ..., and j = 1, 2, ... It should be represented that way for all i and j. Now, let's discuss a numerical example.

Consider an experiment where 3 balls are drawn randomly from an urn containing 2 red, 3 white, and 4 blue balls. Let us consider an experiment of drawing 3 balls randomly from an urn. The urn contains two red balls, three white balls, and four blue balls. There are two red balls, three white balls, and four blue balls. The balls are randomly mixed, and it is equally likely that three balls are drawn.



There are a total of 4 + 3 = 7 + 2 = 9 balls. So, the 3 balls drawn may contain 2 red balls and 1 white ball, or it may contain 3 blue balls. However, it is not possible to have 3 red balls since there are only 2 red balls. It is also possible to draw 3 white balls. These are the possible outcomes when drawing 3 balls.

Now, the question is: Let (X, Y) be a random variable where X and Y represent, respectively, the number of red and white balls chosen. So, X is the random variable representing the number of red balls chosen, and Y is the random variable representing the number of white balls chosen. So, we have a random selection of red and white balls. We need to find the range of X and Y. The first question is: (a) find the range of X and Y. Let's start by finding the range of X.



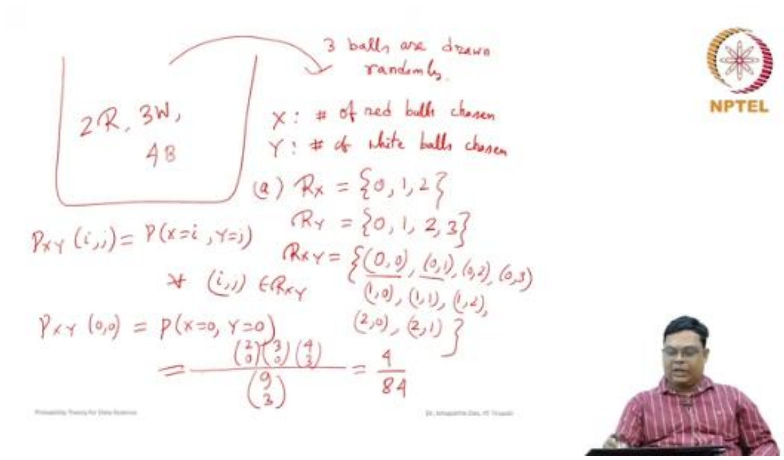
Now, whenever you draw 3 balls, what would be the possible number of red balls? The number of red balls could be 0, which would happen if all 3 balls are blue. It could be 1 or 2, as there are a maximum of 2 red balls. The range of Y will be as follows: it could be 0 white balls if there are 3 blue balls, meaning no white balls are present. However, it is possible to have 1 white ball, 2 white balls, or 3 white balls. Now, let's find the range of (X, Y).

I have just found the range of X and Y. Now, let's find the range of (X, Y). The range of (X, Y) refers to R_{XY}. So, what are the possible values for this bivariate random variable? It will be a subset of R^2. More specifically, it will be a subset of R_1. So, what are the possibilities? For example, the red ball could be 0, and the white ball could also be 0; this is a possibility. The red ball may represent 0, and the white ball may represent 1. Another possibility is that the red ball is 0, and the white ball is 2, or the red ball is 0 and the white ball is 3.

These are possible combinations. Now, the red ball being 1 and the white ball being 0 is also possible. The red ball being 1 and the white ball being 1 is another possibility. The red ball being 1 and the white ball being 2 can also occur. However, the red ball being 1 and the white ball being 3 is not possible. So now, if the red ball is 2 and the white ball is 0, it is possible. If the red ball is 2 and the white ball is 1, it is also possible. In total, there are three possibilities. There are no more possibilities: 1, 2, 3, 4, 5, 6, 7, 8, and 9. Therefore, there are 9 elements in this set within the range of (X, Y).

This defines the range of (X, Y). Now, find the joint probability mass function of X and Y. The joint probability mass function can be determined by finding P(X, Y) for (i, j), where i and j belong to the set R(X, Y). These are the only possibilities that need to be checked. This represents the probability of X equals i and Y equals j. Therefore, we need to find this probability for all (i, j) that belong to (X, Y). I am not writing (xi, yj) explicitly because here xi is equal to i, and yj is equal to j. For simplicity, let us write X equal to i and Y equal to j. Now, we will find the probabilities one by one. For example, suppose we consider this element. This means we want to find P(X, Y) for (0, 0). This is simply the probability that X is equal to 0 and Y is equal to 0. So, X equal to 0 and Y equal to 0. If you want to find the probability, that means this is nothing but the number of cases where, if you draw the three balls, the number of red balls is 0 and the number of white balls is also 0. So, how can it happen?

What will be the possibilities? Now, how can we find the probability? Now, because it is equally likely, the total number of possible ways to draw the balls is calculated. There are 9 balls, and you are choosing 3 balls randomly, which is 9C3. From the red balls, you are choosing 0, so for the red balls, it is 2C0, which gives 3 possibilities. From the white balls, you are also choosing 0, and from the 4 blue balls, you are choosing 3. That is the possibility where the red balls are 0 and the white balls are also 0. Now, if you simplify these, 4C3 is just 4. For 2C0 and 3C0, it's 1. 9C3, as we've computed, is 84. You can check that 9C3 is indeed 84. Similarly, you can find the other values. Suppose we consider (0, 1). This represents the probability P(X, Y) for (0, 1), which is the probability that X is equal to 0 and Y is equal to 1. The total number of possibilities is 9C3.



The number of red balls is 0, so that's 2C0. The number of white balls is 1, so that's 3C1. The remaining 4 balls are selected 2 at a time, so that's 4C2. These are the possibilities. So, the probability is calculated as follows: since i is 0, 1, 2, and 0, we find 18 by 84. You can check if this is correct. This is because 3C1 is 3 and 4C2 is 6, so 6 times 3 is 18, and 18 divided by 84 gives the result. So, the probability is 18 by 84. Similarly, in general, you can write P(X, Y). Now, we can understand how to compute this and write it generally.

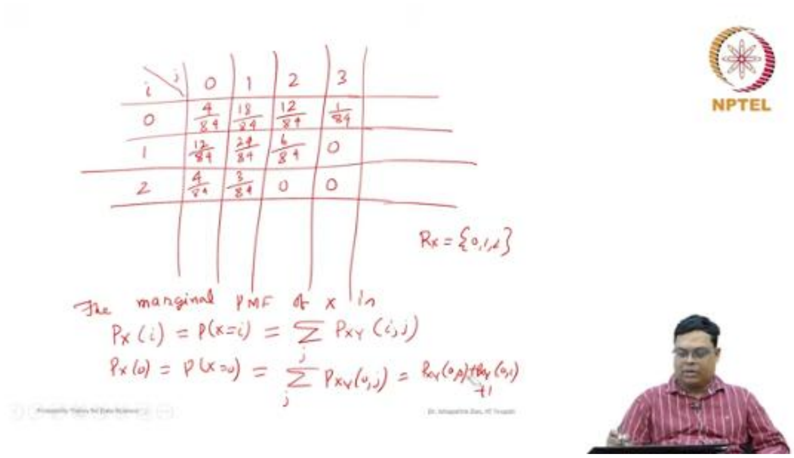
So, in general, we can write P(X, Y) of (i, j), which is simply the probability that X equals i and Y equals j. Then, this probability is calculated by choosing i red balls out of 2, j white balls out of 3, and the remaining balls from 4. This gives the total number of balls, which is 3 minus i plus j. This is the combination, which is 9 choose 3. This equation works when $i \geq 0$ and $i \leq 2$, and $j \geq 0$ and $j \leq 3$. Additionally, $i + j \leq 3$. i + j can also be 0, as in the case where i = 0 and j = 0. For all other cases, the probability is 0. Now, if you represent this graphically as a table, it would look like this: i and j. Since there is limited space, let us write it here. So, here, i is 0, 1, 2, and j is 0, 1, 2, 3.

For (0, 0), we found that the probability is 4 by 84, and for (0, 1), it is 18 by 84. Similarly, you can calculate the other values as well. As shown here, (0, 0) is 4 by 84, (0, 1) is 18 by 84, (0, 2) is 12 by 84, and (0, 3) is 1 by 84. Now, for (1, 0), it is 12 by 84; for (1, 1), it is 24 by 84; for (1, 2), it is 6 by 84, and for (1, 3), it is not possible because the total number of balls is 3. Since 1 + 3 equals 4, the probability is 0. Now, for (2, 0), the probability is 8 by 84; for (2, 1), it is 3 by 84; and for (2, 2), it is not possible, which is why the probability is 0.

So, we have found the joint probability mass function. Now, let's find the marginal probability mass function of X and Y. We will compute the marginal probability mass function. How can we find that? The marginal probability mass function is as follows: 18 by 84, 12 by 84, and 1 by 84. Then, 12 by 84, 1 by 84, and 1 by 84. The next line is 12 by 84, 24 by 84, 6 by 84, and 12 by 84, 24 by 84. Sorry, 24 by 84, then 6 by 84. This is not possible. The next line shows the following values: 4 by 84, 3 by 84, 4 by 84, and 3 by 84.

These are correct. The rest are 0. Now, to find the marginal probability mass function of X, we will calculate P(X) of xi. Here, xi is equal to i, so this is nothing but the probability that X is equal to i. So, because the range of X is 0, 1, and 2, we need to find the probabilities for these values. By definition, we have already determined what this is. We need to take the sum of all other values of P(X, Y) for i and j. For example, if you want to find P(X) of 0, this is simply the sum of all P(X, Y) for j, given that X is equal to 0. So, this is simply

P(X, Y) of (0, 0) plus P(X, Y) of (0, 1) plus P(X, Y)... Let us move on to the next page and find the marginal probability mass function.



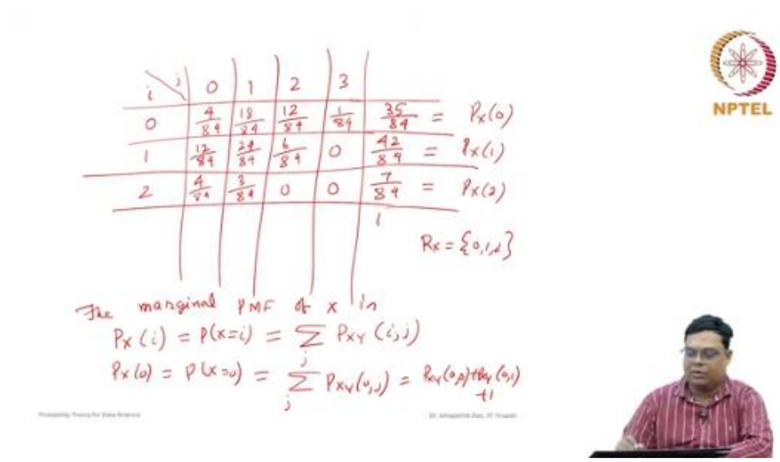
The marginal probability mass function of X is given by P(X) of i. This is simply the probability that X equals i. We know that the formula is Pxy(i, j). Here, i belongs to the range {0, 1, 2}. So, if we want to find P(X) of 0, this represents the probability that X equals 0.

The summation of j can range from 0 to 3, as the possible values of j are {0, 1, 2, 3}. Pxy(0, j) is found by fixing i. This is equal to Pxy(0, 0) + Pxy(0, 1) + Pxy(0, 2) + Pxy(0, 3). By fixing i = 0, we want to find Pxy(0, 0), Pxy(0, 1), Pxy(0, 2), and Pxy(0, 3). This means that if you take the row sum here, you will get P(X) of 0.

Similarly, you can see that if you take the row sum here, you will get P(X) of 1. If you take the row sum here, you will get P(X) of 2. Now, what is the value of P(X) of 0? If you sum 18 + 12, you get 30; 30 + 4 is 34; 34 + 1 is 35. So, the value is 35/84.

This is P(X) of 0. Now, P(X) of 1 will be 12, then 36; 36 + 6 is 42. So, this will be 42/84. This is P(X) of 1. P(X) of 2 will be 7/84. If you add the row sum, this is 7/84.

Now, if you take the sum of all these values, it should match 81. As you can see, 35 + 42 + 7 = 84/84, which is equal to 1. Now, how can we find the marginal probability mass function for this? We got 35/84. Similarly, you can compute the other values. So, now, what will be the marginal probability mass function of Y?



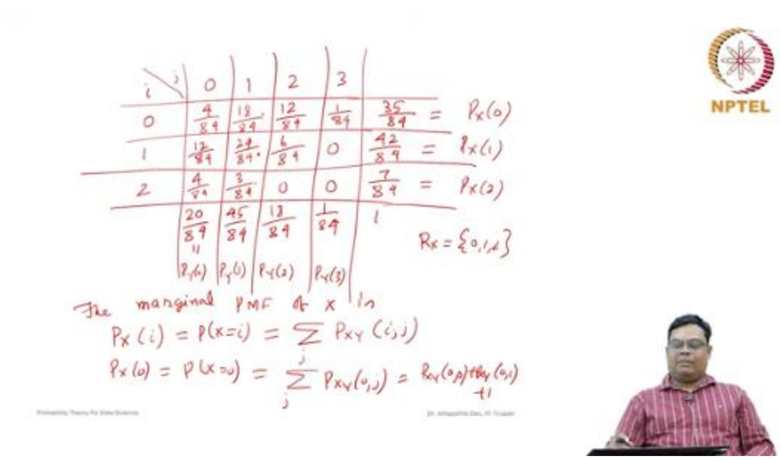
The marginal probability mass function of Y is given by P(yj). This is nothing but the probability that Y equals j. So, we know that if you take the sum of the other values from 0 to 2, X can take Pxy(i, j) for j belonging to {0, 1, 2, 3}, which represents the range of Y. The range of Y is {0, 1, 2, 3}.

Now, suppose you want to find P(Y) of 1. This is the probability that Y equals 1, which is the summation of Pxy(i, 1), where i ranges from 0 to 2. So, it is nothing but Pxy(0, 1) + Pxy(1, 1) + Pxy(2, 1), with i varying and 1 fixed for j. So, here you can see that Pxy(0, 1) is this, and Pxy(1, 1) is this. Then, Pxy(2, 1) is this. So, because we are fixing 1 here, the sum of column 1 gives us P(Y) of 1. This is the same as P(Y) of 1.

Similarly, if you fix 0, the column sum gives us P(Y) of 0. Hopefully, you are following this. So, the row sum gives us the marginal of X, and the column sum gives us the marginal of Y. This is P(Y) of 2, and this is P(Y) of 3.

Now, let's compute the values. This equals 20/84. This is nothing but 18 + 24. So, 24 + 18 equals 42. Then, 42 + 3 equals 45. Next, 45 divided by 84, and this is the same as 18/84. This simplifies to 1/84.

Now, you will take the sum. If it is correct, then the total should match, and when you take the sum here, it should equal 1. In this way, we can find the marginal probability function or probability mass function of Y.



Here, you can find the marginal probability mass function of X. So, the question is to find the marginal probability mass function of X and Y, and we have found that.