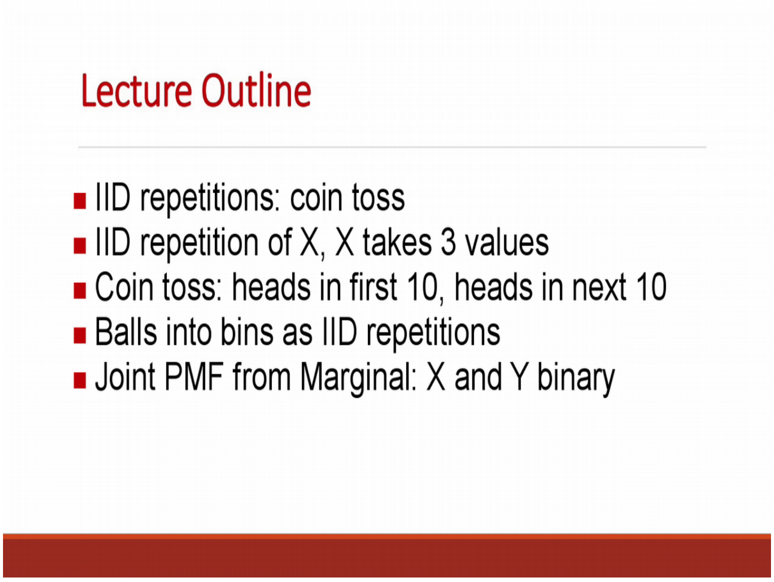
Probability Foundations for Electrical Engineers
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

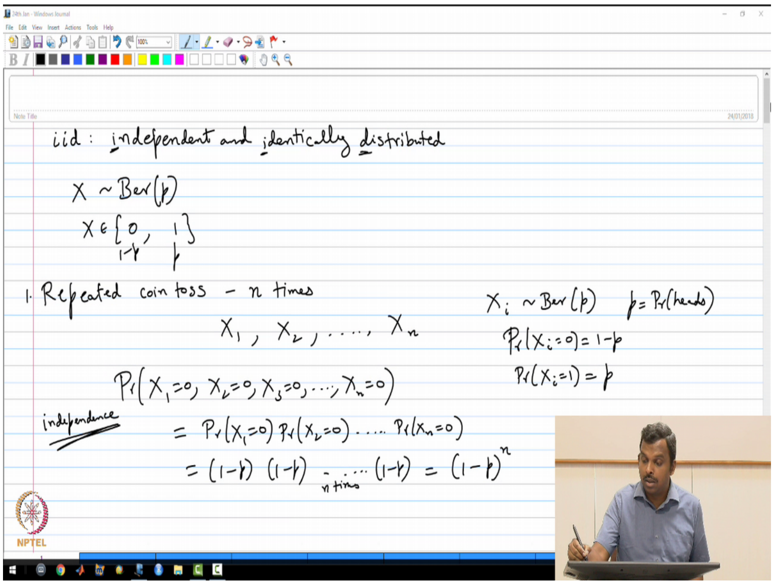
Lecture – 38
Examples: IID Repetitions

(Refer Slide Time: 00:13)



Welcome to this lecture with the continuing with examples in probability. We are going to deal with independence and dependence today. In particular in the first video I am going to focus on IID repetitions of particular experiment.

(Refer Slide Time: 00:45)



So, what happens when you have an IID sequence? This word IID is very very important will show up again and again and probability in various context. So, IID means independent and identically distributed, this is where this IID comes in.

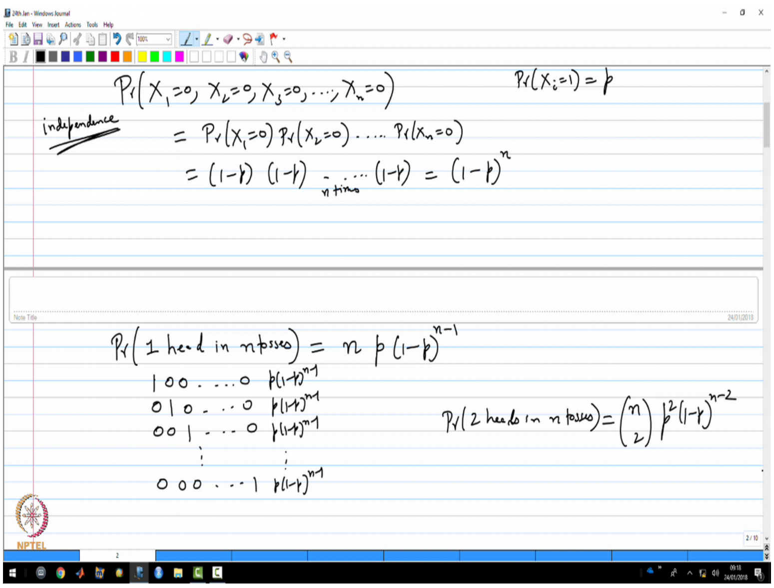
So, let us start with a very simple example. Let us say I have a X being Bernoulli, X is Bernoulli with p. So, what is the meaning of this when you say X is Bernoulli p? The random variable X takes 2 values 0 and 1 value 0 probability 1 minus p and the value 1 with probability p. So, that is the meaning of this, p n 1 minus p right.

So, now, so this is like a coin toss right. So, you toss a coin once you get a Bernoulli random variable. So, 1 could be the heads 0 could be tails. So, that is the way in which you can think of this is an outcome of a coin toss. So, quite often you might have repeated coin tosses right. So, first example we will see is repeated coin toss where you will have Bernoulli random variables coming again and again and again. So, what do I have? I have a fair coin which I toss let us say I toss n times. So, I will have n outcomes and each of these outcomes I will denote by a random variable. So, X 1 is the outcome of the first toss X 2 is the outcome of the second toss so on till X n. So, each of these is the outcome of the ith toss. So, X i is Bernoulli p. So, set p here, p is the probability of heads when you toss the coin. So, you can think of p as 1 half. So, it is a fair coin otherwise it can be any other value.

So, these are the n outcomes. And we have said that these n experiments are independent in the sense that every experiment has nothing to do with other. So, if you have a situation like that you can put together probabilities by simply multiplying. So, that is the main idea. So, for instance this is you X 1 to X n. So, when I say X i is Bernoulli probability that X i is 0 is 1 minus p probability is that X i is 1. So, now, if I have to ask the question probability that X 1 is 0, X 2 is 0, X 3 is 0, so on till everything is 0. So, when you have independent and identically distributed random variables and asks the question what is the probability that all of them are 0s, 0 0 0 0 0 is because they are IID, because they are independent you will have first this happening. Remember once again very very clearly this is because of independence.

So, in case independence is not true which the successive coin tosses are not independent this is always not guaranteed. So, this is true only when they are independent. And on top of that they are all identical. So, they all have the same probabilities. So, you have 1 minus p times 1 minus p so on till 1 minus p n times right, so this happening n times. So, this probability is 1 minus p power n. So, it is an easy calculation to do when you have independent and identical repetition of the things happening. So, for instance you might have ask more complicated questions. So, you can ask slightly more complicated questions like we have done before, what is the probability that there is 1 head in n tosses.

(Refer Slide Time: 04:51)

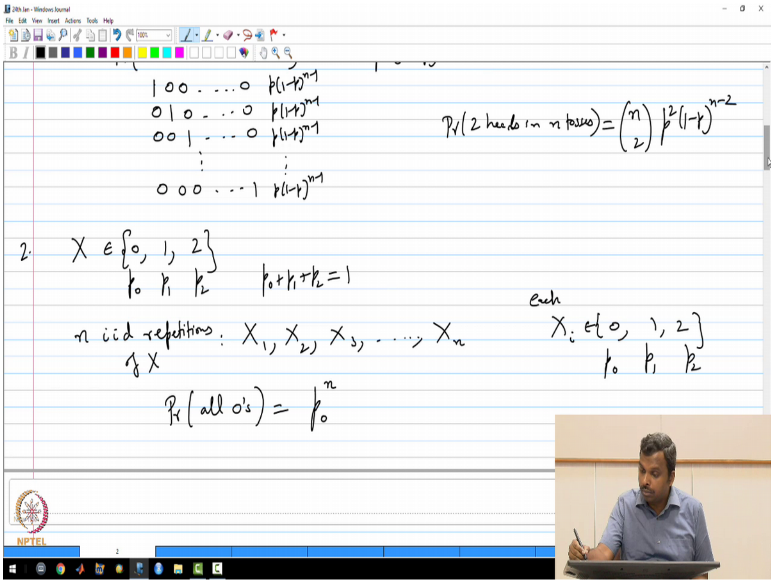


Exactly 1 head and it could occur in the first position or it could occur in the second position or it could occur in the third position and so on and if you look at it there are n possible ways in which the head can occur and probability of head is p, probability of tails is 1 minus p and that happens and minus 1 tosses. This is the binomial event and you can see the multiplication occurring here right p into 1 minus p raise to the power n minus 1, there n minus 1 tails, n minus 1 0's and only 1 head. So, 1 single 1 occurs in the sequence X 1 to X n. So, for that 1 you have a probability p and for all the 0's you have probability 1 minus p. You multiply them together you will get n minus 1 guys. So, you have 1 minus p power n minus 1 and whereas, the n come from this head p can occur either in the first position or in the second position has n possible places. So, you have to add up all the possibilities.

So in fact, if you want to write it down this corresponds to the event 1 0 0 0 0 0 or 0 1 0 0 0 0 or 0 0 1 0 so on till 0 0 0 1 each of these guys has probability p into 1 minus p power n minus 1 right. So, each of these guys has this probability you add up all of them together you will get n times p 1 minus p power n minus 1. So, this is how you deal with IID events. It is easy to count these cases sometimes the calculation can get a bit painful. So, for instance if you want the question probability of 2 heads n n tosses. So, here you will you will have a slightly more complicated answer n choose 2 p square 1 minus p power n minus 2. So, you will see p squared into 1 minus p power n minus 2 is easy to see, but then how many possibilities are there with exactly 2 heads that is n choose 2. So, this is how independent events are dealt with.

You can have slightly more complicated situations which we will look at next in the second example.

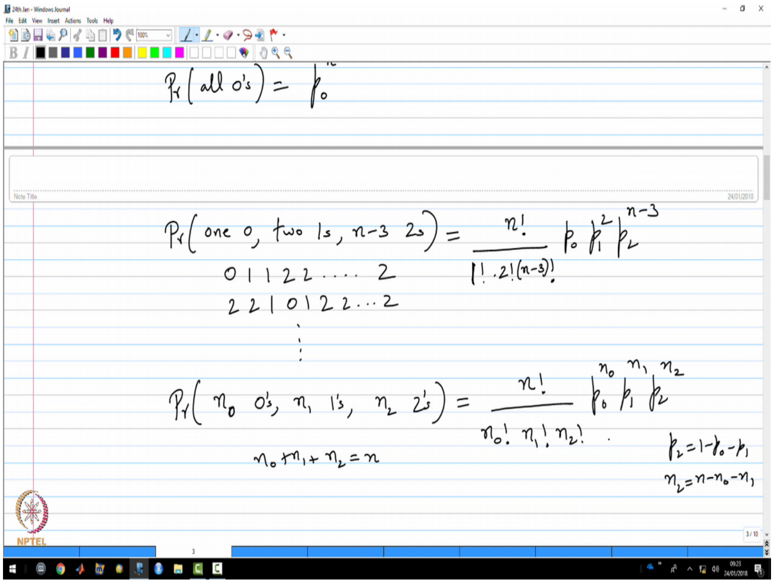
(Refer Slide Time: 07:10)



So, I am going to say X is random variable it takes let us say 3 values 0, 1, 2 and the probabilities are p 0, p 1, p 2 and remember p 0 plus p 1 plus p 2 has to be equal to 1 right. So, this has to be valid probability distributions. So, I am looking at a random variable which instead of taking 2 values 0 and 1, now takes 3 values 0 1 and 2. So, now, I am going to look at n iid repetitions of X and I will call them X 1, X 2, X 3 so on till X n. So, what does that mean? Each X i takes 3 possible values 0 1 2, values 0 with probability p 0, value 1 with probability p 1, value 2 with probability p 2. So, now, so remember again these are all iid repetitions of this random variable. So, each random variable here takes 3 values and that can happen n times.

So, now, again you can ask such question. So, what is the probability that you have all 0s? So, all of them are 0's once again you can multiply all of these case together you will get this very simple answer p 0 power n.

(Refer Slide Time: 08:39)



So, if you want to ask slightly more complicated question. So, supposing you say probability of let us say 1 0. So, let me write it down clearly 1 0, 2 1's, and n minus 3 2's right. So, what is the meaning of this? So, I have in my sequence X 1 through X n exactly 1 of these values is 0 could be anywhere, exactly 1 of these values is 0 2 of the values are 1, so 1 0 2 1's and then all the remaining of 2's. So, that is the kind of event I am considering for instance it might be 0 1 1 2 2 2 2 2 or it can also be 2 2 1 0 1 2 2 2 like this. Like this you will have so many events like that lot of possibilities here know where does 0 come, where does the 1 come and where does the 2 come all such possibilities are there.

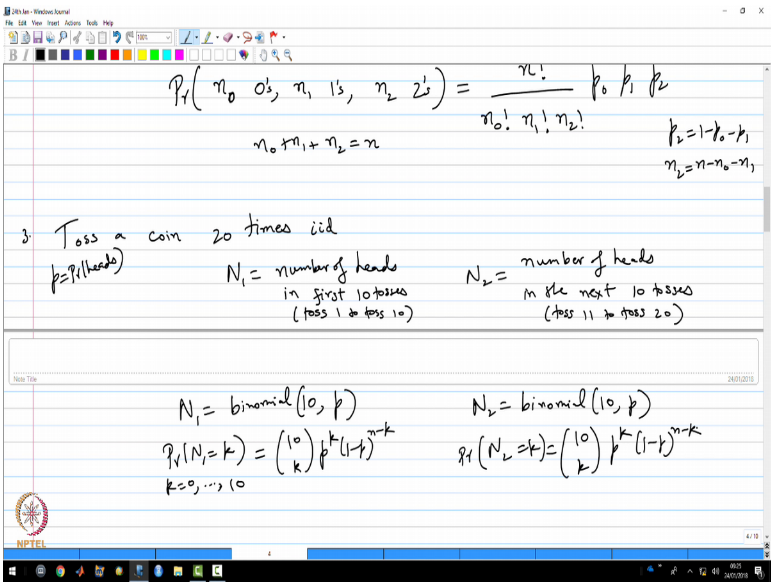
So, turns out if you do the calculation carefully 1 term which is easy to write down is the product of the probability see we will have p 0 p 1 squared p 2 power n minus 3 right. This is easy to write because there is 1 0 2 1's then n minus 3 2's. How many such sequences are there? So, if you count it out carefully you will get n factorial by 1 factorial times 2 factorial times n minus 3 factorial. So, this is the generalization of the binomial formula it is the multinomial term that you get. So, how many possibilities are there with so many 1's, so many 2's, so many 0's, so many 1's and so many 3's. So, that is 1 factorial, 2 factorial, and n minus 3 factorial. So, this is the answer.

So, if you want me to generalize and write it out very carefully its if you have probability of n 0 0's, n 1 1's and then n 2 2's remember and n 0 plus n 1 plus n 2 will be actually equal to n there are n of them n 0 of them are 0's, n 1 of them are 1's, n 2 of them are 2's. So, if you do this general formula will be n factorial n 0 factorial n 1 factorial n 2 factorial p 0 power n 0, p 1 power n 1, p 2 power n 2. So, this is a generalization of the binomial formula. So, we had before when you have n tosses and you ask what is the probability of k 1's you write n factorial by k factorial into n minus k factorial times p power k 1 minus p power n minus k this is a generalization of that.

Remember n 2 is n minus n 0 minus n 1 and similarly p 2 p 2 is also 1 minus p 0 minus p 1 right. So, remember that. And 2 is n minus n 0 minus n 1. So, this is fact that which should be satisfied. So, it is you can use this formula to add probabilities in when you have IID repetitions. Excuse me. You can use this formula to add probabilities when you have IID repetitions now you can also generalize this. I am not going to talk about it here, but you can its possible.

So, now, there is this slightly different example I want to talk about.

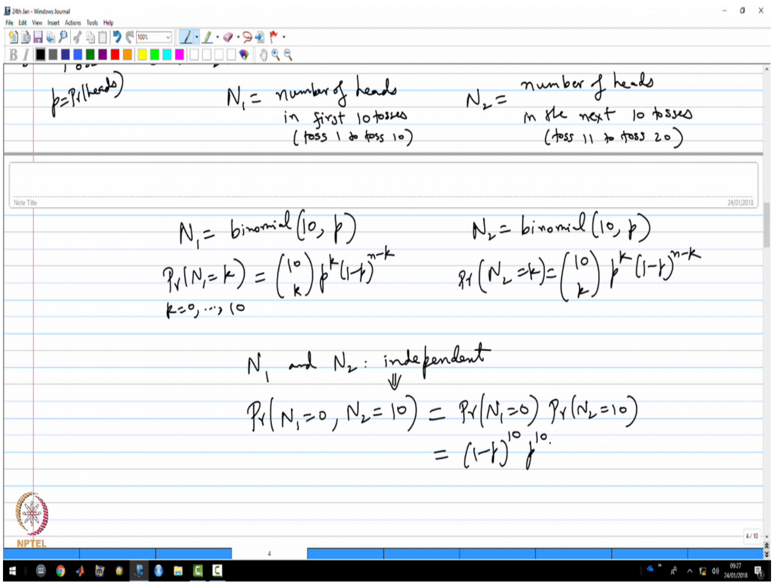
(Refer Slide Time: 12:05)



Suppose you toss a coin 20 times. So, I am gone toss a coin 20 times and I will define N 1 as the number of heads, remember this is also going to be iid we will we will make that assumption in first 10 tosses. I will say N 2 is the number of heads in the next 10 tosses. So, what do I mean by this when I say first 10 tosses? Toss 1 to toss 10 and when I say next tosses I mean toss 11 to toss 20. So, this is the thing, I am defining I am tossing the coin 20 times the first 10 tosses how many heads I got then next 10 tosses how many heads I heads I got. So, that is the 2 random variables N 1 and N 2.

So, N 1 itself is easy to write down and one is going to be binomial with these 2 parameters n comma p. So, p is going to be the probability of heads of a single toss. So, similarly N 2 will also be I am sorry here, so this 10. N 2 will also be binomial with 10 n p right. So, what do I mean by that? So, probability that N 1 is k equals 0 to 10 is actually 10 choose k p power k 1 minus p power n minus k same thing holds here probability that N 2 equals k its again the same thing n choose k p power k 1 minus p power n minus k. But remember n 1 is a function of what happened in the first 10 tosses and n 2 is the function of what happened in the next 10 tosses and this two are not connected in any way because they are all from independent tosses right. So, as a result what happens is N 1 is independent of N 2. So, that is the important result here which we will use.

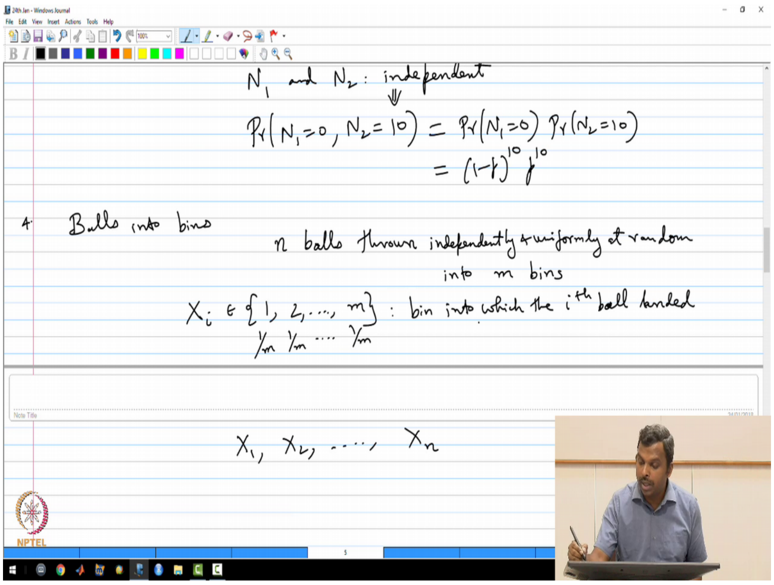
(Refer Slide Time: 14:26)



So, important observation is that N 1 and N 2 are independent because did not overlap in anyway right. So, N 1 is a function of what happened in the first n tosses n 2 is the function of what happened next in next 10 tosses. So, these 2 also will be independent. So, when you have derived some functions with independent random events you will also have random variables being independent. So, N 1 and N 2 are independent. So, if somebody were to ask you a question what is the probability that N 1 is 0 and N 2 is 10 one can multiply, because they are independent this implies one can multiply these things. So, we can do probability of N 1 equals 0 times probability of N 2 equals 2.

So that means, if they are not independent you cannot do this multiplication. So, because they are independent one can do this multiplication. So, this is easy to do. So, probability that N 1 is 0 1 minus p power 10 probability that N 2 is 10 that is p power 10 indicate a very nice answer of this form. So, if they were dependent then you want to do more complicated calculations, but for independent IID guys since they were functions of independent events this works out very easily. So, let me do a one more example of how this IID repetition comes in and how it useful.

(Refer Slide Time: 15:50)



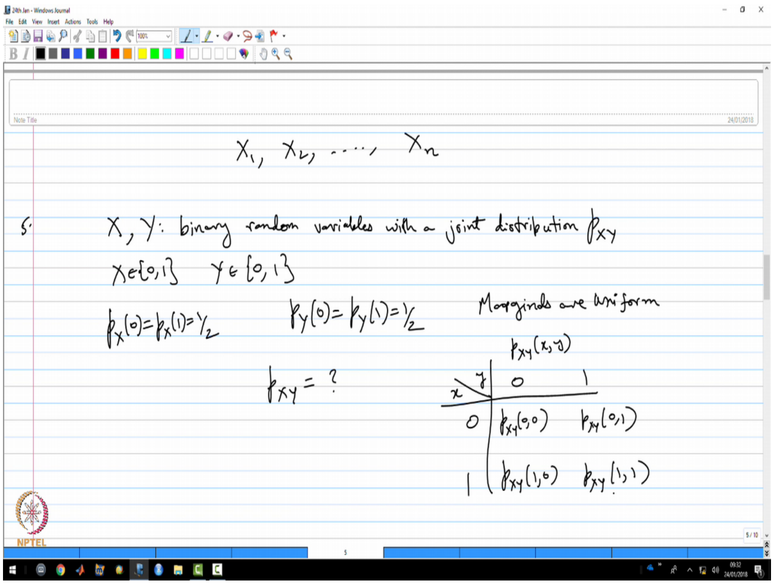
So, we will be talking about this balls into bins experiment a lot. So, let us say we consider n balls thrown a independently and uniformly, uniformly at random into m bins right. So, this is the experiment been considering a lot.

And remember each ball what you do with each ball is uniform it can, ball can go into any bin and it is also independent of what you did with the previous ball or the next ball or anything like that. So, there is no connection between these things. So, that is very important. So, one can think of this that is an IID repetition. So, let us say I define X i to take a m values 1 2 to m and I will say this uniformly distributed. So, what is this X i? This is the bin into which the ith ball landed.

So, now, I have a sequence of random variables X 1, X 2, X n with X i denoting the bin into which the ith ball landed. So, this is an IID repetition and each X i is uniform and takes m values 1 to m each with probability 1 by m and I have an IID repetition. So, this balls into bins like it is like an IID repetition in this sense. So, in this sense if you looking at the number of the bin in which the ball landed and the end of the day that is of interest you that is your outcome you can think of it as an IID repetition. So, you can use this to do simplify some of the calculations in balls in to bins it is not always very easy, but you can use this to simplify some of the calculations it will help you. So, this is way in which IID repetitions can happen in simple experiments coin toss as one examples even balls into bins falls under that category of examples of IID repetition in some sense.

So, the last example I want to do is this quite important and it will enforce the idea of independence and dependence and all of that ok.

(Refer Slide Time: 18:14)

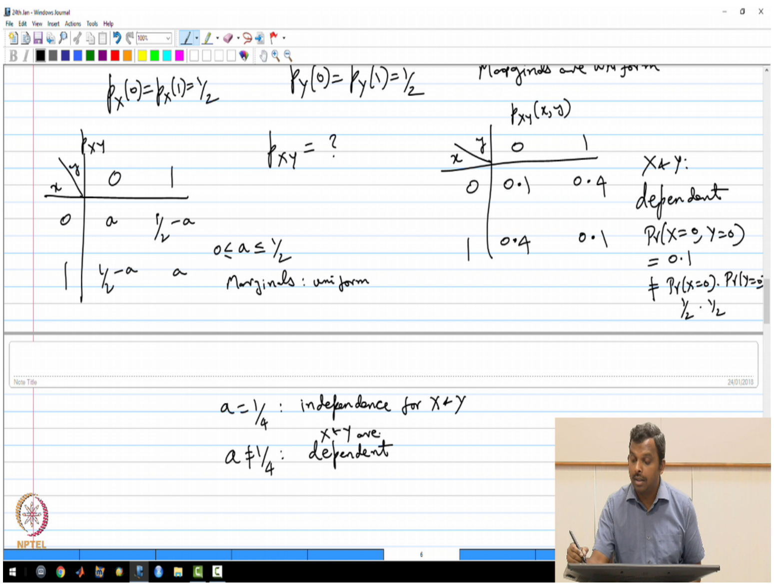


So, I am going to consider that case where X and Y are binary random variables with a joint distribution p XY. So, there is some joint distribution for X and Y they are binary random variable. So, what do I mean by binary? X takes 2 values 0 1, Y takes 2 values 0 1. So, I am going to say I know the marginals of these 2 random variables what do I mean by marginals I know p X p X of 0 and p X of 1 it says uniform. I know this is half and likewise I know p Y of 0 and p Y of 1 are both half.

So, marginals are known marginals are uniform right what can you say about the joint distribution that is the question. So, the marginals of these 2 random variables they jointly distributed I know that and then I have their marginal distribution I know that X itself is uniform Y itself is uniform by itself. What can we say about the joint distribution p XY? Now a lot of people would they immediately say X and Y is independent. So, p XY is actually going to be one-forth. So, what do I mean by p XY? There are 4 values right. So, you have, it is convenient or write it like this in the form of a table. So, you have a X here, Y here, X is 0 1, 0 1 and this table captures p XY.

For instance this could be p XY of 0 comma 0 this p XY of 0 comma 1 now I should be careful here I will put X here and Y here, so p XY of 1 comma 0 and p XY 1 comma 1. So, table like this captures the joint PMF. So, I have set the marginals this half. So, lot of people would simply multiply half with half and put with 1 by 4 here and 1 by 4 here, 1 by 4 here, 1 by 4 here. So, that is just one possible joint PMF which gives you these marginals there can be many other joint PMFs. So, for instance you can write something like this right.

(Refer Slide Time: 20:44)



So, I could have something like 0.1 here, 0.4 here, 0.4 here, 0.1 here. So, this you can check is a valid joint PMF you add up all of them you get 1 and your marginals are still 0.5, 0.5, 0.5. So, how do you get the marginal of Y? You have to add these 2 things 0.1 plus 0.4 plus 0.5. So, marginal for Y equals 1 its 0.4 plus 0.1 again 0.5. So, every marginal is 0.5. So, X and Y are still uniform, but the joint distribution is not the independent case right.

So, for independent case I should get 1 by 4 here, 1 by 4 here, 1 by 4 here, 1 by 4 here, but this is not that, this is some other joint PMF for which you have X and Y being uniform. So, this is a dependent situation. So, X and Y are dependent. So, you can check why that is so, because probability that X equal 0 comma Y equal 0 right, this is equal to 0.1 and it is not equal to probability of X equal to 0 times probability of Y equals 0. So, this is not true. So, because of that reason you have this being dependent you can see why that is true, now because this is 1 by 2 into 1 by 2 this is 0.25. On the other hand this is just 0.1. So, this is a joint distribution which is dependent.

In fact, this is not the only joint distribution for which this is true you can have any other joint distribution, so you could take for instance another joint distribution, XY being 0 1 0 1. I can take this value to be some a and I will put half minus a here, put half minus half here and a here and a could be any value in the range 0 to, I could take it from 0 to half. So, for any a from 0 to half this joint PMF will always give you marginals which are uniform for both X and Y.

So, how many joint distributions are there which give you marginal uniform PMF for 2 for binary value at random variables? Actually an infinite number, it can be any number. You take any a from 0 to half you will get in fact, there is an extreme example I can give you where it will still be a marginal in that fashion. For instance you could take a equal 0. If you take a equal 0 you have 0 0 half half. So, it is very sort of an extreme. So, X and Y together take only 2 values. If X is 0 Y is 1, if X is 1 Y is 0 that is all. So, it is kind of deterministic relationship between X and Y or you could take a equals half in which case X and Y are equal, X is 0, Y is 0 or X is 1, Y is 1 which still set joint the marginal distribution is uniform or the joint distribution is anything, but independent. What is the only independent case? If a is equal to 1 by 4 you have independence. So, remember that, a equals 1 by 4 you have independence for X and Y. So, for a not equal to 1 by 4 X and Y are dependent.

So, this is something an important lesson to just to re-enforce with the notion of independence and where it comes from. You just from the marginals, you cannot guess anything about the dependence or independence, we need the joint distribution and there can be any number of joint distributions which give you the same marginals and the independence scales is just one possibility. So, hopefully you got an idea of how to use independence, what is the meaning of IID repetition and all of that in this lecture.

Thank you.