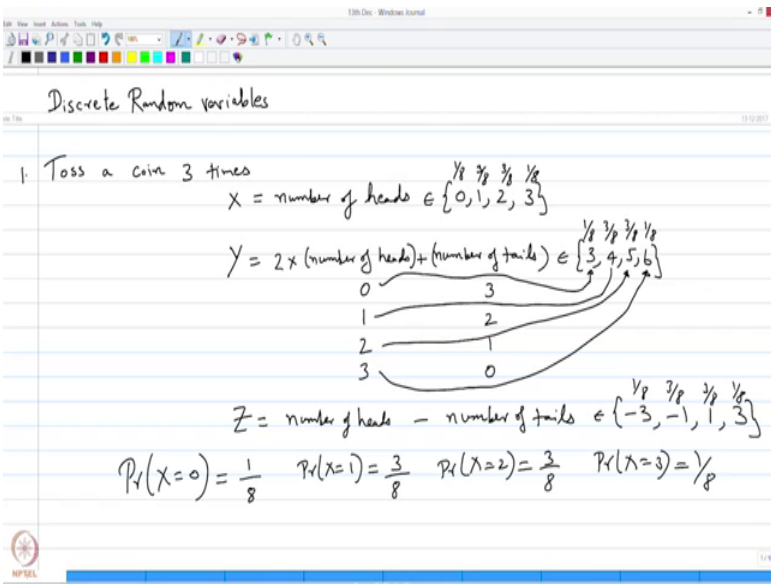
Probability Foundations for Electrical Engineers
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 23
Examples: Discrete Random Variables

In this speak you must have started studying about random variables this is actually one of the most important things that you need to really know what is the random variable, what is the definition of it, what is the how to calculate things with the random variables. So, I am going to start examples, I am going to continue with my examples similar to before tossing a coin, throwing a die, cards and balls and bins etcetera then start defining random variables and talk about them. You will see in one way it is easy, but it is also very important and you will see why this random variable idea is quite interesting and let us get started.

So, what we are going to see examples for, just like Professor Aravind is talking about, mostly we will talk about discrete random variables at this point.

(Refer Slide Time: 00:58)



So, the first examples I am going to start I will start again with a toss a coin let us say 3 times just like before. So, you have an experiment and then you can start defining random variables based on the outcomes of the experiment. So, random variables the thing is they should take numerical values.

So, of course, what is a numerical? Supposing you have a set which is heads comma tails you can call heads as 1 and tails as 2 and you again have a numerical situation right. So, it is possible to do that, but usually people would have real life experiments whether outcomes may be the outcome naturally itself is a number, but even if it is not a number you can define some numbers and then look at them. So, that is what I am going to do I am going to define some random variables which are numbers based on the sub serration for instance if you have an experiment like toss a coin 3 times. I could have a random variable X which denotes the number of heads. So, this could be the random variable.

The first thing to understand when somebody defines a random variable is what are the values that it can take in the experiment? So, this is always the first thing to write down. So, we know that the number of heads in this experiment if you toss the coin 3 times could either be 0 or 1 or 2 or 3. For discrete random variables this will be a set which you can write down in this fashion. So, it is 0 comma 1 comma 2 comma 3 you can easily write down all the elements in the set for a discrete random variable. So, this is the number of heads. So, I could define various other random variables for instance, I could define a random variable Y which is let us say you know I mean you can come up with. So, many different definitions for instance 2 times the number of heads plus the number of tails let us say. So, this if you actually carefully look at the possibilities this will actually take 3 4 5 or 6 this will be the number of values that it will take think about it.

The number of heads could be 0 in which case the number of tails would be 3 the number of heads could be 1, in which case the number of tails could be 2, number of heads could be 2 in which case the number of tails would be one and the number of heads could be 3 in which case the number of head tails would be 0. So, in an experiment these are the possible outcomes and then if you plug it in you will either get 3 or 4 or 5 or 6 those are the 4 possible values. So, like this you can easily convince yourself that you can define as many number of random variables that you like. So, for instance you might want to define a random variable Z which is number of heads minus number of tails right.

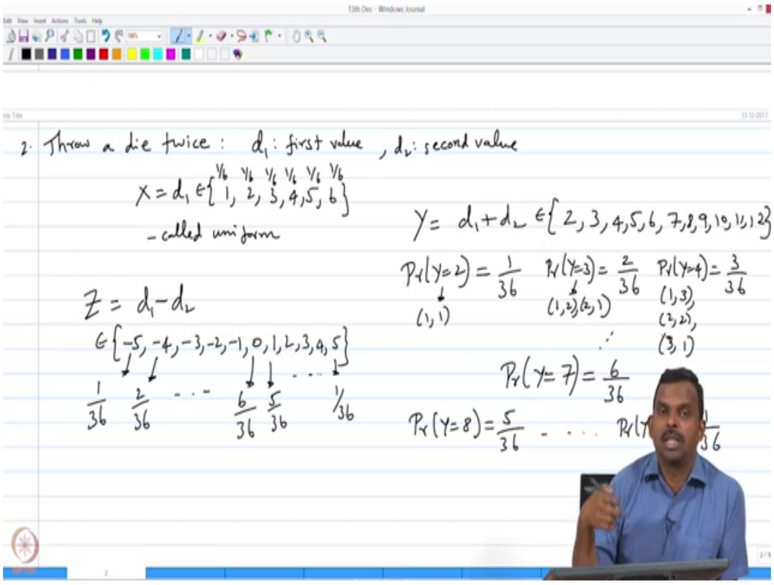
So, what are the kind of values that it could take? It can take minus 3 it can take the value minus 1, it can take the value plus 1, it can take the value plus 3 do you agree again go back to this table number of heads could be 0 in which case number of tails would have become 3 number of heads could be one in which case number of tails would be 2 number of heads could be in which case number of tails would be one etcetera. So, it is say the 0 minus 3 or 1 minus 2 or 2 minus 1 and 3 minus 0. So, that is the possibility here is that hopefully it is clear to you.

So, these are all I mean simple random variables that are defined. So, usually the most important thing to do when you have a random variable is you have to first find out what are the values that the random variable can take and what is the probability that it takes that particular value. These are the 2 things that one needs to compute when number one defines a random variable. So, for instance here I defined the random variable a number of heads which takes 4 possible values 0 1 2 3, the next thing I might be interested in is what is the probability that X equals 0. So, this calculation results in what is called the PMF of probability mass function of the random variable it is very very important to be able to compute that or know that.

So, what is the probability that you have 0 heads that is 1 out of 8, what is the probability that X equals 1 you have one head that is 3 possibilities 3 out of 8, what is the probability that X equals 2 that is again 3 out of 8 and what is the probability that you have 3 heads and that is 1 out of 8. So, one of the things I like to do when I write down a random variable is I would write what the values it takes and I will write the PMF value right on top of the value that it takes. So, I calculated the probability of 0 is 1 by 8, probability of 1 is 3 by 8, probability of 2 is 3 by 8, probability of 3 is 1 by 8. So, that is a complete description of the random variable. To describe the random variable you have to say what values it takes and what is the probability with which it takes those values and this is how you do it.

So, what about the random variable Y? So, it is 2 times the number of heads plus number of tails, but really you know the first value of 3 happens when the number of heads is 0. So, again this is just 1 by 8, 3 by 8, 3 by 8, 1 by 8 do you agree think about that. So, this corresponds to number of heads being 3 being 0; sorry and this corresponds to number of heads being 1, this corresponds to number of heads being 2 and the last guy corresponds to number of heads being 3 right. So, that is just 3 1 by 8, 3 by 8, 3 by 8, 1 by 8. And the same thing here, so what is this guy this is again probability. 1 by 8, 3 by 8, 3 by 8, 1 by 8, these easy to do for these kind of experiments if you define random variables like this. Hopefully this was clear to you. Let us move on to the next experiment which is through a die twice.

(Refer Slide Time: 07:15)



So, I am going to define 2 random variables here you throw a die twice let us say d 1 is the first value and d 2 is the second value. So, we threw the die twice we have got 2 values first value was d 1 second value was d 2. Now, d 1 could be any one of 6 possibilities 1 2 3 4 5 6, d 2 could also be any one of 6 possibilities 1 2 3 4 5 6. So, I can define a random variable let us say X equals d 1. So, it is a very simple random variable it would take 6 values and what will be the probability which I can write on top 1 by 6, 1 by 6, 1 by 6, 1 by 6, 1 by 6, 1 by 6. So, such random variables are called uniform.

You can contrast with the previous case the previous case our random variables warrant uniform they took what they could take 4 values and the probability with which they took these values are different, but here you have this random variable which takes 6 values and the probability that it takes any one value and they are all equal 1 by 6 if you have a case like that it is called uniform. Now, I can define slightly more complicated random variables I might want to define a random variable which is let us say d 1 plus d 2 the sum of the 2 values. Now what are the values that this can take? It can take 2 3 right 4 5 6 7 8 9 10 11 12, could be anyway from 2 to 12 and you can start think about trying to compute this probability what is the probability that Y equals 2 the only case in which the sum can be 2 is if you got one comma one right. So, that is 1 by 36. So, what is the probability that Y equals 3 you can get 3 if you got 1 comma 2 or 2 comma 1 right. So, this is 1 comma 1, so it is 1 by 36, this could be 1 comma 2 or 2 comma 1 which is 2 by 36.

So, what about probability that Y equals 4, here you will get 1 comma 3, 2 comma 2, 3 comma 1, so that was going to be 3 by 36 is that. Hopefully you can see that this number will keep on increasing right. So, if I ask for the general value supposing something like probability that Y equals say all the way up to 7. So, if you look at 7 you are going to have a 6 by 36 right. So, 7 will be 1 comma 5, 2 comma 4, I am sorry 1 comma 6, 2 comma 5, 3 comma 4, 4 comma 3, 5 comma 2 and then 6 comma 1. So, the first toss can be any 1 of 6 possibilities and then the other toss will be fixed by the first. So, you get 6 by 36.

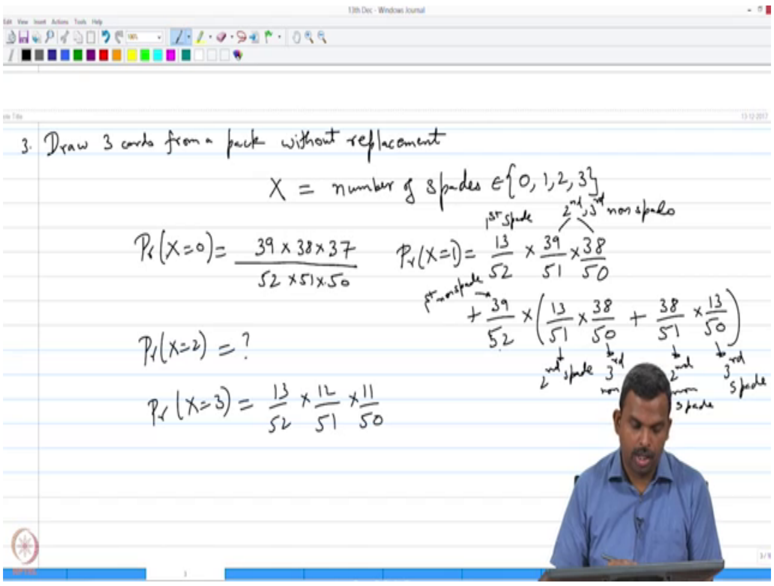
So, now, what happens if you go? So, likewise it will go 1 by 36, 2 by 36, 3 by 36 all the way to 6 by 36 and then you will have probability of Y equals 8. Now Y equals 8 what will happen? It will actually reduce it will become 5 by 36 why because all 6 possibilities are not there for the first one if the first toss resulted in a 1, you can never have a sum of 8 right. So, the first toss should be at least 2. So, you have only 2 to 6 for the first toss. So, I have 5 by 36 likewise all the way till probability of Y equals twelve and that will become 1 by 36 again versus 6 comma 6. So, this is a way to compute these probabilities.

You might also want to define a random variable Z which is maybe d 1 minus d 2. So, what will happen if you do a d 1 minus d 2? This is a little bit more intriguing you might get you might have minus 5 right minus 4 minus 3 minus 2 minus 1, 0 1 2 3 4 5 these are the possibilities. In fact, if you think about it very very carefully you will have a very similar distribution to the sum for the difference also. In fact, what will be the probability of this. So, if it has to be minus 5 the first toast has to be one the second toss has to be 6. So, that is just 1 out of 36 possibilities what about minus 4 you can have 1 and 5 or 2 and 4. So, that is 2 by 36 all the way till 0 you will actually have 6 by 36. Well you have can have 1 comma 1, 2 comma 2, 3 comma 3, 4 comma 4, 5 comma 5, 6 comma 6, all 6 possibilities and then it will start falling down this will be 5 by 36 and then all the way to 1 by 36. So, this is the kind of distribution you will get.

So, hopefully you can see defining random variables based on the numbers that you get in your experiment is quite straightforward computing the probabilities can be a little bit tricky. But if you write it down you can do it this once again when somebody defines a random variable in an experiment you have to first find out what are the values it takes and then try to find what are the probabilities with which it takes those values and it can be easy or complicated based on the problem.

So, I am now going to move on to slightly more complicated experiments where the maybe these calculations are a little bit more tricky.

(Refer Slide Time: 12:53)



So, let us look at draw 3 cards from a pack without replacement. So, you have a pack of fifty 2 cards once again I am going to draw 3 cards a well shuffled pack and then you pick 3 cards at random and what are the kind of random variables that one can define. I am not replacing. So, if you do not do replacement you remember you have to use some conditional probability arguments and carefully count situations to find the probabilities you will see that will come up in these cases also, but let us see what is the random variable we can define.

So, I am going to just look at one statistic one random variable here let us say X is the number of spades. So, now, clearly if you have 3 cards from a pack there are 12 13 spades in the suit of cards. So, you can in the pack of cards. So, you can have either 0 spades or 1 spade or 2 spades or 3 spades all 4 are possible. So, this seems straightforward enough the value that the random variable takes is just 4 possibilities.

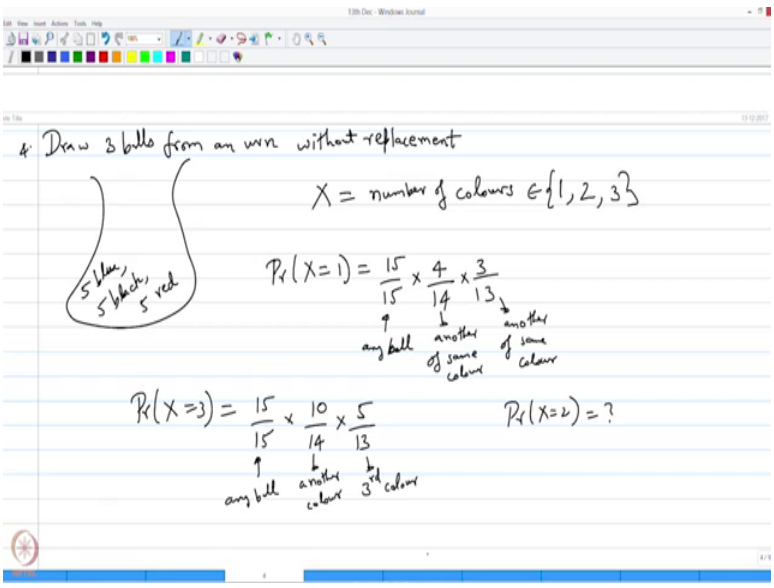
Now, what about the probability? So, here when since you are doing it without replacement one needs to be a little bit careful. So, if you have 0 spades. So, you have to you have to count carefully right. So, the first card should not be a spade that is actually one of 39 possibilities. So, first I think it is good to write down the denominator what are the total number of possibilities. First card could be any 1 of 52, next card could be 51, next card could be 50 right. So, this is a total number of possibilities now we have to find the possibilities when the how many possibilities are there when you do not have any spades. So, if you do not have any spades the first case is 39 because there are 39 cards which are not spades and then once you have drawn a not spade what you have left is thirteen spades and 38 non spades and you have to pick from the thirty 8 and likewise what is remaining finally, will be 37 this is the probability that X equals 0. So, this is again a case of conditional probability you can condition on the first event and so on.

Now, what about probability that the card X equals 1. So, you have one spade. So, here again you have to split based on what happens in the, remember this is exactly one spade. So, if I say one it is exactly one spade I should have 1 spade and 2 non spades. So, I might have the case that the first card I picked was a spade. So, this is the probability that the first card I picked is a spade. So, if I pick a spade in the first card what should happen in the remaining 2 draw? I should not pick a spade. So, how many non spades are left 39 out of 51 and then 38 out of 51. So, that was the first possibility when I picked a spade or I may not pick a spade, I may not pick a spade in the next possibility.

Now, after this there are 2 possibilities again my second card could be a spade which is 13 by 51 probability and if my second card is a spade and then my third card should not be a spade right. So, that will be 38 by 50, I am sorry I put a 51 here it should be 50, 50. Once again remember the first is spade second and third are non spades. So, here I have first being non spade, here I have second spade and here I have third being non spade or you could have the second being a non spade which is 38 by 51 in which case in which case the third has to be a spade. So, this is a way to compute probabilities this is the conditioning method. You conditioned on what happens in the sequence of events and this is something that you have to do.

So, likewise you can do for probability of X equals 2, I will leave that as a homework you can think about it what could you do. I will do the easier case of probability that X equals 3 in this case all 3 of spades. So, the first one has to be a spade, the next one has to be a spade, the third one also has to be a spade there are thirteen spades to start with out of 52. Once you have taken a spade out there are only 12 spades left. So, 12 out of 51 remaining cards and then there are 11 spades left 11 out of 50 remaining cards that is the probability case. Once again probability that X equal to its a exercise that little involve something like this your condition on the whether or not the first is a spade and then the second could be a spade or a non spade and based on that you will get different possibilities the first is a non spade second and third has to be a spade. So, you should just write down what each thing happens in a sequence and then since we are doing without replacement independent draws you can multiply the probabilities and condition carefully will get your answer. So, this is the drawing 3 cards from a pack without replacement.

(Refer Slide Time: 18:05)



So, I am going to do another problem like this which is slightly complicated draw 3 balls from an urn and the urn has let us say 5 blue, 5 black and 5 red, balls it has 5 blue balls, 5 black balls and 5 red balls. I am drawing 3 balls from the urn without replacement. So, the without replacement makes everything a little bit complicated or interesting depending on your point of view, but without replacement is the case we are going to consider and then once again just look at one random variable this is the number of colors that you got you drew 3 balls. All 3 could be of the same color in which case the number of colors would take the value 1 or you could have 2 balls of one color and another ball of a different color in which case the number of colors is 2 you could have say 2 blue balls and 1 black ball.

So, you have 2 different colors or you could have all 3 colors is that, you could have either all 3 of the same color remember that could be all 3 blues all 3 blacks or all 3 reds or you can have 2 of 1 color and 1 of the other color. So, it could be 2 blue 1 black, 2 black 1 blue, 2 blue 1 red, 2 red 1 blue like that or you could have all 3 of the same all 3 being different colors, so 1 blue, 1 black 1 red.

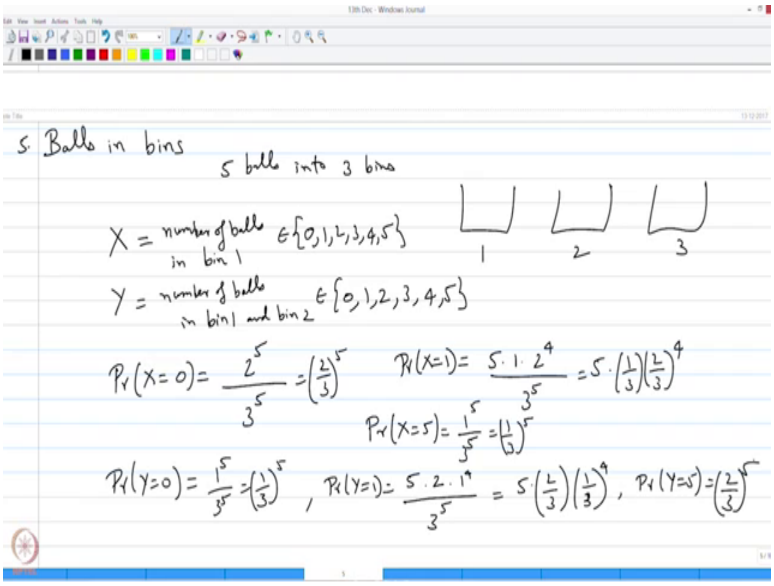
Remember I am drawing this without replacement it could be in any sequence. So, it is a little bit confusing to count this properly. So, let us see. So, this is the values it takes are quite easy now when you want to find the probabilities you have to pay some attention and take some care. It is not too hard, but one needs to be careful. So, let us say if you want to find probability that X equals 1. So, you want only one color all 3 should be of the same color, but remember the first draw can be anything.

It could be a blue or a black or a red I really do not have a condition on the first block right. So, it could be anything, but once I picked a particular color ball I want only one color. So, the next thing should only be that same color. So, now, how many balls of the same color are left once I have picked one ball you have only 4 out of the remaining 14. So, pick 2 balls of one color and the third again becomes 3 out of 13 is that. So, think about this first could be any color any ball another of same color same color as the first ball right and another of same color. So, this is what is most important in this calculation you should you should use conditioning and the fact that the different drawers are actually independent, but you are done without replacement. So, the situation changes a little bit when you condition on what happens is the first one things can change based on that.

So, now, what about probability that X equals 3. So, I want 3 different balls of 3 different colors once again the first ball can be of any color there is really no restriction on the first one. So, once I picked a ball of one color what should happen in the second ball it should be some other color how many balls of other color are left 10 out of 14 and then I have picked 2 balls they are all of 2 different colors now the third ball should be the third color how many are left there 5 by 13. So, this is how you compute this once again what is the logic here, this could be any ball the second of another color, this is third color. So, hopefully you see the logic in this how conditioning works and how the independence works.

Now once again I am going to leave probability that X equals 2 as a homework for you it is one needs to pay attention to the cases the number of cases will be a little bit more tricky here because you want 2 colors, but it is possible to enumerate it carefully and then consider all the cases and you will get it. So, this is the 4th example.

(Refer Slide Time: 22:26)



So, I am going to look at the last one which is balls in bins. So, this I will take a fairly simple example here. So, let us say we throw 5 balls into 3 bins, 1 2 3. So, you can define. So, many random variables and they are all quite interesting. So, you can say a number of balls in bin 1 maybe I will define Y as number of balls in bin 1 and bin 2.

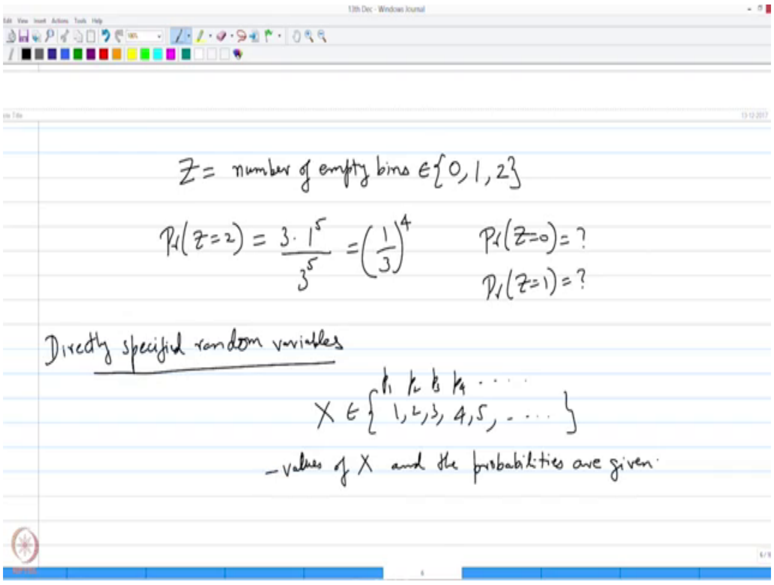
So, now, number of balls in bin 1 takes 0 1 2 3 4 5, number of balls in bin 1 and bin 2 together also takes 0 1 2 3 4 5. Do you agree think about it. So, a little bit I mean it is not too difficult, but logic is quite simple. So, number of balls in bin 1 takes values 0 1 2 3 4 5, number of balls in bin 1 and bin 2 together, so together if I consider what is the probability that it is going to be 0 1 2 3 4 5 etcetera, is that all right.

So, now, computing probabilities here it is going to be a bit interesting. So, let us say probability that X equals 0 is what I am interested. Remember each of these balls could go into 1 of 3 bins we have 3 power 5 in the denominator I want none of them to go into the first bin right. So, all of them should go into either the second bin or the third bin. So, you have 2 power 5 possibilities. So, another way to write it is 2 by 3 power 5 now what about probability that X equals 1 that exactly 1 ball and 2 bin 1. Once again denominator is 3 power 5. What about the numerator? 1 ball should go into bin 1 and that could be any 1 of the 5 balls and then that has only one possibility the remaining 4 have 4 possibilities. So, this is again written in some fashion 5 into 1 by 3 into 2 by 3 power 4. In fact, you can do calculations like this. So, the other pop you can make you can keep doing for the other cases also later on we will see a more general case by which you can do these things, but for now let us keep it like this.

So, if you want X equals 5 all 5 of them to be falling into the same bin. So, you have just 1 power 5 by 3 power 5 it becomes 1 by 3 power 5 think about that that is the probability here. So, what about probability that Y equals 0? None of the ball should be in bin 1 or bin 2 which means all the ball should go to pin 3. So, that is one power 5 by 3 power 5 right, so 1 by 3 power 5. That is just one possibility all of them going to bin 3. So, what about probability that Y equals 1 for instance, you have 5 balls exactly 1 ball should go to either bin 1 or bin 2 and all the other should go to bin 3. So, the build ball that goes into bin 1 or bin 2 could be any 1 of the 5 balls times 2, 2 possibilities for that one ball right it could either go into bin 1 or bin 2 and then the remaining 4 balls have to go into bin 3. So, this divided by 3 power 5. So, you will have 5 times 2 by 3 times 1 by 3 power 5 power 4 I am sorry.

So, likewise if you calculate actually probability of Y equals 5 you are going to get 2 by 3 of 5. So, one can do random variables like this. So, it turns out in the balls on bins question to continue a little bit here.

(Refer Slide Time: 26:23)



So, if you define a random variable Z, but there can be some random variables which are actually quite easy to describe in many experiments, but slightly painful to compute probabilities. For instance I could define Z to be number of empty bins. How many bins can be emptied I am throwing 5 balls in to 3 bins, what is the probability that you have empty bins, how many empty bins will you have that is the kind of question we are asking here.

Now, it is easy to say what values this Z will take it could be 0 there could be no empty bins that is possible or there could be exactly one empty bin, the other 2 are occupied only one bin is empty or you could have what 2 empty bins and that is it you cannot have 3 empty bins right. So, 1 bin at least well have all the balls that will happen. So, you can have exactly 2 empty bins exactly one empty bin or no empty bins so in fact, this is easy to write down. But now if you want to find probability that it takes these values this probability is the easiest. So, you have 2 empty bins which means all the balls went into exactly 1 bin right. So, that 1 bin can be chosen in 3 ways, but once you choose that bin all the balls should go there and this is out of 3 power 5 possibilities. So, this is just 1 by 3 whole power 4.

So, how do you compute the other probabilities probability that there is exactly one empty bin? So, these 2 are a little bit more nuanced, it is not very immediate I am not going to do it in this worked out examples, but think about it, I think it is a nice problem to think about how to do these calculations and computations and what are the various possibilities and how you do the calculation.

So, hopefully you saw quite a few interesting examples of random variables, but interestingly quite often in many problems people would know the question would directly define a random variable. So, this is what Professor Aravind calls in his lectures as directly specified random variables. So, instead of defining an experiment and defining a random variable and asking you to compute the PMF quite often a random variable X will be given to take some values. So, maybe a set of values let us say 1 2 3 4 5 so on and somebody will tell you that the probability of 1 is p 1, probability of 2 is p with p 2, p 3, p 4 and so on. So values of X and the probabilities are given. So, this is like just directly specifying, this is the random variable I am considering. And a lot of probability theory involves calculations starting from this point. So, there are two points in which you start, you start with the experiment in a sample space and define a random variable and compute the PMF of that random variables and what is the values it takes and what is the probability with which it takes these values. That is also very important, quite often that plays an important role.

Another important thing set of things to understand is suppose somebody gives you the random variable what do you do from that point, are that things you can do with random variable, study their property study their manipulations etcetera from this point on in the in the classes Professor Aravind we will talk a lot about this. He will start with the random variable and its different distribution and it is a PMF sort to speak and then you start talking about manipulations. So, that part is also equally important to understand. So, I am going to stop here for today I will continue in the next lecture with more examples.

Thank you very much.