Data Analysis and Decision Making - I

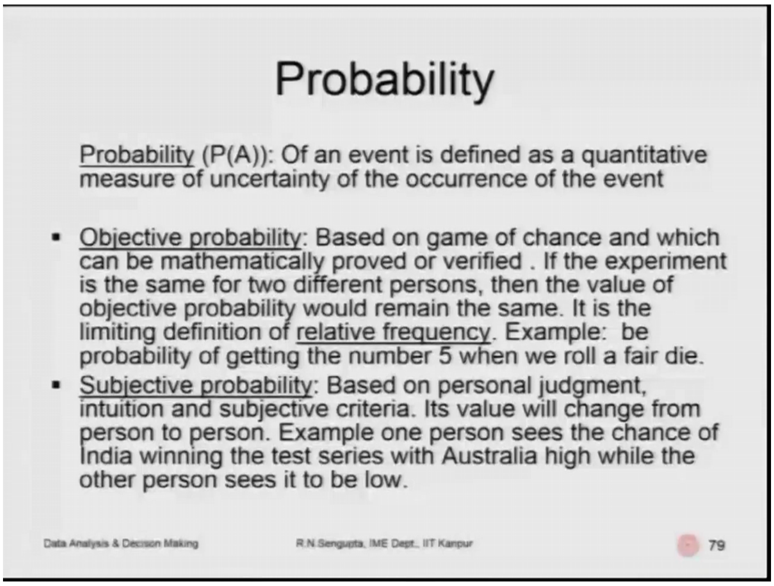
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 05
Introduction to Probability

A very good morning, good evening, good afternoon to all my dear students and friends, welcome back to the 5th lecture under the NPTELs MOOC series and the course is DADM which is as you can see is Data Analysis and Decision Making one which basically is going to deal everything related to statistics and related topics. And as you know, this is the 5th lecture which is would be the end of the first week, we have 12 week course; that is 30 hours lecture each week, we have half an hour and each week would have basically 5 lectures.

So, if you remember you are discussing about probability, the concepts of probability, and how the concept of probability can basically be extended to understand different type of probability distribution functions probability mass function. I will come to those topics in more details.

(Refer Slide Time: 01:05)



So, probability of any event; so, what do you mean by event? What we mean by an experiment and what we mean by probability and where the actual concept of probability comes from considering its actually con related to frequency and relative frequency we will discuss that. So, probability of an event is defined as a quantitative measure of uncertainty of the occurrence of the event. So, we say for example, if we have a dice and all the faces are marked from 1 to 6 and the probabilities or the chance are the chance of other frequency in the long run for any one faces coming is equal and that is they are unbiased.

So, we will say the probability of getting number 1 or 2 or 3 or 4 or 5 or 6 when you roll the die when you playing the game of Ludo is 1 is 1 by 6. Or say for example, if you have a coin; unbiased coin and if you toss it the relative frequency or the frequency in the long run or the probability of trying to get the either the head or the tail or if the faces are marked as blue or red or black or red whatever it is it is half.

Now, whenever you are dealing with probability, you should think of the concept that probability is basically the relative frequency in the long run. So, consider that you have you are doing an experiment consider the very simple one; you are tossing the coin and you do it 100 times. So, say for example, then it is unbiased coin and the number of faces which is coming as head is say for example, 55 and number of faces which is coming as tail is 45.

Now, again you consider. So, on obviously, in that case, the frequencies are as given as 55, 45 and the corresponding relative frequencies are 55 by 145 by 100 which is 0.55, 0.45. Now you do the next experiment again and you again toss it at 100 and consider now the heads and tails in the respective numbers is now 60 and 40; that means, the relative frequency is 0.6 which is 60 by 100 and it is 0.4, this is 40 by 100.

Now, you keep repeating it. So, the next time, the corresponding relative frequency of head and tail is 70 30.7 and then for tail is 0.3 corresponding to next time when you do it is basically 0.5, 0.5 other, again, you do it say for example, is 0.5, 1.49. Next time, you do is say for example, is 0.4 and 0.6 and consider it continues. So, what actually you will get in the long run, if you basically keep repeating it, then the averages of these averages; average means basically the relief frequency which you have is in the first case 0.55, 0.45 or 0.4, 0.6 or say for example, 0.5, 1.49. So, in this case, if you basically keep adding them and trying to find out the averages of their average in the long run for the probability for the head or the relative frequency with the head would come out to be 0.5 as similarly, the relative frequency of the probability for tail would also come as 0.5.

So, objective probability based on the game of chance and which can be any mathematically proved or verified we can do that if the experiment is the same for two different persons. Then the value of the objective probability would remain the same, it is limiting is basically the limiting definition or relative frequency as I mentioned. That means, if we keep repeating the experiment to the relative frequency.

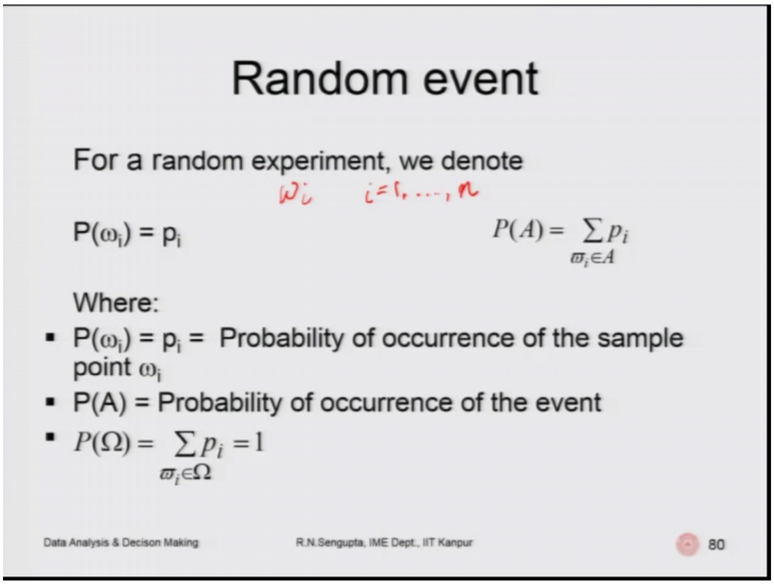
In the long run as the actual number of experiments which you are doing goes to infinity that value is basically the probability example like finding of the probability in getting the number 5 when you roll the fair die in the long run, say for example, the actual probability would be total number of 5s which have coming divided by the total number of such throws you are doing should actually 10 towards 1 by 6, but in any case if you do it, say for example, you roll it 600 times.

So, it need not be that in each and every such experiment which you do the probability or the relative frequency of getting number 5 in each case would be 100 by 2; 600 which 1 by 6, it may differ, but in the long run the relative frequency or the actual average of the average should 10 to basically 1 by 6 subjective probability on the other hand is definition based on personal judgment intuition and subjective criteria it is the value which will change from person to person.

So, example one person sees the chance of India winning the test series with Australia high while the other person will say that the chances of India winning the test match again Australia is average or another person may say its low. So, basically is subjectivity is the point where the individual info set of information in which individual judgment would basically be becoming in a very big way in order to basically have some information about subjective probability.

But in the case of objective probability, it would be a very well defined mathematical formula based on which you are going to study and understand what is probability.

(Refer Slide Time: 05:55)



For an a random experiment so; that means, when you are doing an experiment with the outcome, before you get the outcome, it is unknown, but once the outcome is available you know it, but obviously, that would be among the set of such possible outcomes which is possible for a random experiment will define that each outcome is basically a sample point.

So, say for example, when you roll the die the sample points are 1 or 2 or 3 or 4 or 5 or 6, now if I am define an experiment that is the that is the experiment is that when you get the even number only then the collection of the sample points which is 2; 4 and 6 would basically give you the information number about the experiment, if I say that I want to do an experiment whether number is less than equal to 4. Then the sample points would be the point number 1 or plus 2 plus 3 plus 4, because it is less than equal to 4. So, whenever you are saying in a sample means sample points would be defined by omega I; where i is basically equal to the number of such the sample points which you have it can be infinite also remember that I will come to the experiment or the examples later on, and we will denote that the probability of getting that sample point is given by small piece of xi depending on the case that what does I denote.

So, it say for example, if I am rolling the die and the probabilities are equal; that means, it is an unbiased die. In that case the probability of getting number one is P 1 which is equal to 1 by 6 probability of getting w 2 which is number two is equal to P P suffix 2 which is 1 by 6 probability of getting in the number 3 where w 3 is equal to 3 is equal to P suffix 3 small P suffix 3 is equal to basically 1 by 6.

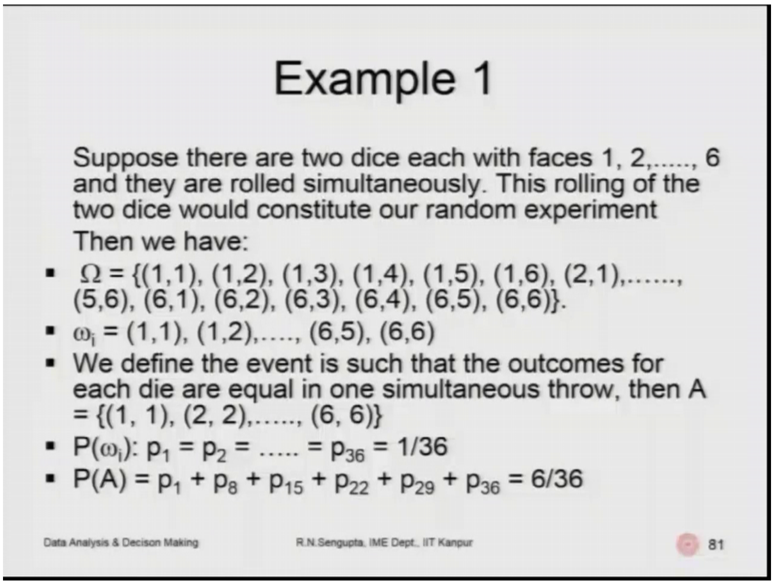
So, in this case it will continue, but in case say for example, if I say that the corresponding probabilities when you are tossing the coin the corresponding probability of getting a head is say for example: 2 by 3 and the probability of getting a tail is 1 by 3 in that case when w one is head the probability P 1 is equal to 2 by 3 when w 2 is equal to tail the probability is equal to basically is equal to 1 by 3 which is P 2 P means caps small P suffix 2.

So, depending on the experiment which you have you will analyze and say then the collection on the point sample points which basically gives you the experiment a. We are using a symbol a to denote the experiment is basically what the actual probability for that event is. So, say for example, we are saying for an unbiased die if you roll the coin if you want to find out that what is the probability of getting an even number then the w w's corresponding to that event would be w 2, w 4, w 6 and the corresponding probabilities are 1 by 6, 1 by 6, 1 by 6.

Then the corresponding overall probability for that event will be 1 by 6 plus 1 by 6 plus 1 by 6. So, here we will denote the capital P w i which is small p i is the probability of occurrence on the sample point w i and the probability of a is the probability of the occurrence of the event where the event is given by the conglomeration of all the sample points which are corresponding to that event a. So, the event a can be can have basically infinite set of points will come to that later points and; obviously, if you take all the sample points which makes the universal set.

So obviously, it will mean the corresponding probability of the universal set would be one. And on the other hand with the corresponding probability of the null set would be 0. Suppose there are 2 dice with phases, 1 to 6 and they rolled simultaneously.

(Refer Slide Time: 09:33)



One after the other or at the same time this rolling on the 2 dice could constitute our total random experiment, how it is let me denote specify then. So, in that case the overall sample space or the total universal set would be in one case consider what are the outcomes the outcomes can be the first one is 1. That means, of the die and the second one can be either from 1 to 6 in this in the other case it can be the first one is 2 and the outcomes for the second one is 1 to 6; similarly for the third case, fourth case, fifth case and 6 case.

So, if I considered all of the different type of combinations of the sample points your actual outcome would be as denoted first one would be one comma one second one would be 1 comma 2. So, hence forth till the last one where we will have 6 comma 5 which where the outcome for the first die 6 and the outcome of the second die is 5 and the last one will be the outcome of first die and the second die is 6 comma 6, we define the event in such that the outcome for each die are equal in into 1. In one assumption is throw, then the experiment would be given or the event a would be given where we have one to basically find out what the experiment is that if the outcomes for both the dice are same. So, it will be 1 1, then it will be 2 2, then it will be 3 3 and it will be 4 4, then the 5 5 and 6 6.

So,; obviously, the corresponding probabilities would be 1 by 6 plus 1 by 6 1, 1 by 6 plus 1 by 6 plus 1 by 6 1 by 6. So, we are basically trying to 1 by sorry 1 by 36 because my apology is 1 by 36 because the outcomes are 1 and 1. So obviously, the outcome of getting a one in the first die would be 1 by 6 and the getting a number 1 in the second dice 1 by 6. So, it will be 1 by 36, similarly getting 2 2 would be 1 by 36. Similarly getting 3 3 would be 1 by 36, similarly 4 4, 4 5, 5 6 is each of them would be 1 by 36. Hence, the probability of getting the event where the numbers are equal on the dice which are rolled would basically be we will be counting 1 by 36; 6 number of times.

So, if you check here the first value P 1 means 1 1 in the first and the second die the 8, 1 would basically mean we have 2, in the first and the second die 3, 3 in the first and the second 4, 4 in the first and second 5, 5 in the first and second 6, 6 in the first and second. So, hence the total corresponding probability 6 by 36 suppose the coin, we start tossed repeatedly till the first head is obtained.

So, in this case you will understand what I have been saying that the overall number of such sample points can be infinite. Now if you do sorry. So, if you note down the experiment you want to do it tose or toss a coin repeatedly in the first head is obtained. So, in this case what the outcomes can be it is in a head and you stop the experiment in the other case if the if the first outcome is tail then you will continue consider you get ahead in the second throw, then you stop it.

So, if it does not happen in the second through what is the output what is the next possible sequence the possible sequence would be you get a tail in the first, you continue, you get a tail in the second, you continue to get a head in the third and you stop. Then if you do not get in a head in the third what are the possible sequence, it will be tail for the first tail for the second tail for the third head for the fourth stop. If it does not stop, then it will be they would be tail head tail, one tail, second tail, third tail, forth tail, fifth head, stop; so in this overall sequence. Obviously, would have such infinite sequences hence the overall set on the number of sample points which you will have for the experiment would be infinite.

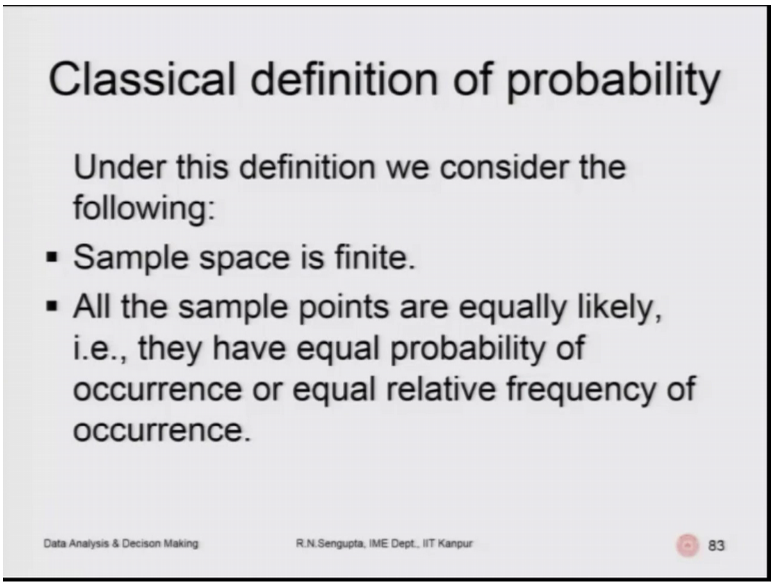
So, let me denote it how what I mean talking about. So, in this case the universal set would be the first sample point would be a head you stop it second sample point would be a tail and a head third one would basically be a tail, tail head fourth one would be 3 tails and a head and so on and so forth. So, if the sample points are a head as I said a tail and head. So, a tail, tail, head or tail, tail, tail, head and continues now if we define the event such that we want to basically have at most 3 tosses and needed to obtain the first head.

And also consider that: what is the probability of getting each and every sample points. So, the sample point where you get a head in the first trial is basically half and the sample point where you get a head in the second trial; and a tail in the first trial would be half into half the sample point where you get a head in the third trial and a tail and a tail in the first two would be half into half into half. Similarly, for the fourth case will be half to the power 4 then half to the power 5 half to the power 6 and so, on and so forth, it will continue.

So, in that case the probability of the sample points are p 1 which is half p 2 is half to the power square p 3 is half to the power q p for half to the power 4 and continues. So, in this case if you want to find out that at most 3 tosses are needed to obtain the first head then; obviously, the explain would be a head at most 3 tosses. So, it would be head it would be a tail head and the last case would sample one would be tail head. So, in that case the corresponding probability of the overall event would be for p 1 it is half p 2 is half to the power 4 p 3 is half to the power three.

So, you will add all of them and the value comes out to be 7 by 8.

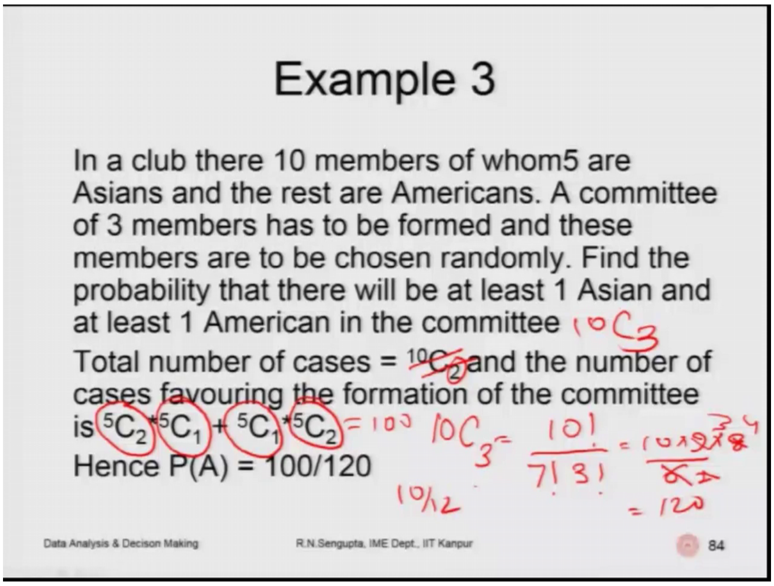
(Refer Slide Time: 15:17)



So, the classical definitions of probability under this definition we consider the following the sample spaces are finite. That means, remember that; that means, that the number of sample points which constitute the worlds sample space is finite all the sample points are of equally likely. That means, the probabilities are same that they have equal probability of occurrences or equal relative frequency are occurrences. Hence trying to find out of the values for the event or the values for the overall sample space is very straightforward.

So, in the case when you want to find out that when you are rolling the dice and what are the probabilities of getting a 1 or a 2 or a 3 or a 4 or a 5 or 6 is all are equal probabilities 1 by 6 the sample points are finite. And hence, finding out the overall sample space or an event is very straightforward. So, this is the classical definition of probability consider this in a club.

(Refer Slide Time: 16:10)

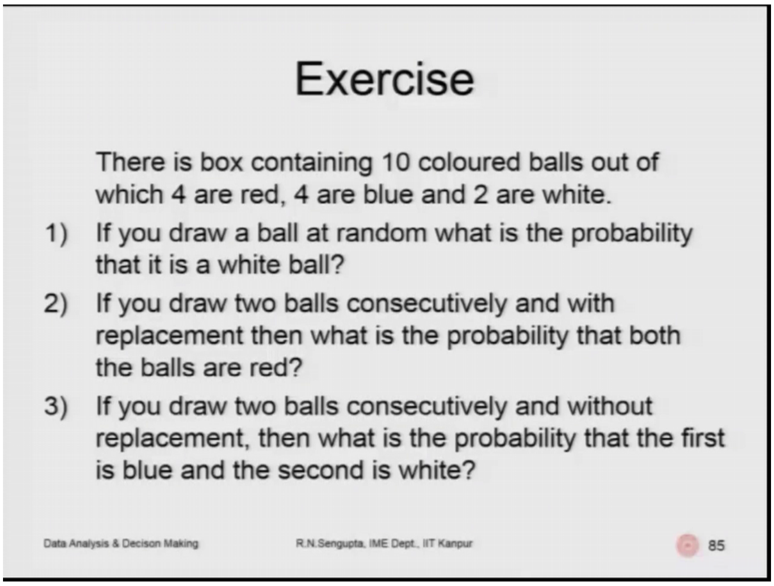


There are 10 members of whom 5 are Asians and the rest are Americans, a committee of 3 members has to be formed and these members are to be chosen randomly find the probability that they would be at least 1 Asian and at least 1 American in the committee. So, there are 1 Asians and 1 Americans, and you want to basically have 3 members; it can be 2 Asians 1 American or 1 Asian and 2 Americans. So, the total number of cases you will have is basically 10 C 2 and the number of this cases where you want to basically if have the combination would be in given as this.

So, the pre check. So, it would be you want to check 3 1. So obviously, it would be 3 yes just give me one minute. So, this will be my mistake this would be 10 C 3 because you have 3 different positions to filled up and the total combined our number of combinations which you can do would be 10 C 3. So, that will come in the denominator and the number of cases actually favoring that committee would go work is 1 Asian, 2 Americans or 2 Asians, 1 Americans. So, what would the combination would be if there are 5-5 members each.

So, in this case there are 2 Asians, 1 Americans and the next case is 1 Asian and 2 Americans. So, the total corresponding probability would be in this value comes out to be hundred. So, 100 divided by 10 C 3 which is 120. So, would the value comes out to be 10 by 12 consider the next experiment, there is a box containing 10 colored balls out of which 4 are red 4 are blue and 2 are white.

(Refer Slide Time: 18:11)



So, if you draw a ball at random what is the probability that is a white ball? So obviously, if you are doing it for the first time and this concept of replacement or non replacement does not come then the probability of getting a white ball would be you have two whites the total number of balls which is basically 4 plus 4 which is basically a plus 2 is 10. So, you will find out the corresponding probability and then basically conclude the experiment.

So, finding of the of the trying finding of the chi then actual number of frequency of getting a white ball would be that you will have either you will pick the first white or you can pick in up the second one also as white. That means, there are two balls, you would pick up any 1 of them. Now if you draw two walls consecutively and with replacement then what is the probability that both the balls are red; so, if you pick up two balls and the first one is red. So obviously, 4 are red and the total number of balls you can pick up is 10.

So, the corresponding probability can be found out, but remember one thing if you draw it consecutively with replacement it key means that you have drawn the first one as red note it down kept into the box. So, the number of red remains 4 number of total balls remains 10. So obviously, to corresponding probability in the second draw also remains the same, but now if I ask the third question if you draw two balls consecutively and without replacement then what is the probability that the first is blue.

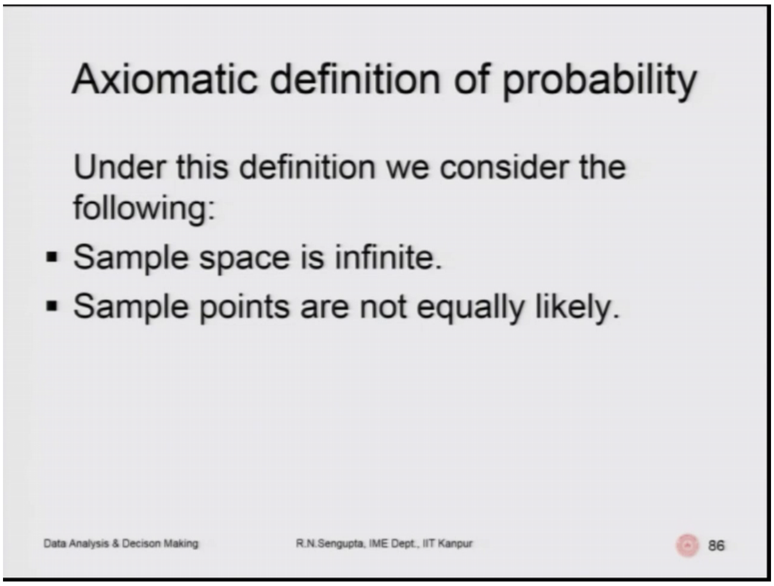
And the second is white or say for example, continuing the same way as we did for the second example what is the probability of would drawing two balls consecutively it red if you draw the first ball and do not replace that in that case number of balls which you have for number of such red balls. You will have on the first trial would they would be two balls out of the 10, then when you pick up that red ball find it out and then replace it. That means, we will have only one ball to be picked up which is red, but also you should note that the total number of such pickings you can do from the total number of balls is not now 10, it is basically 9.

So; obviously, the corresponding probability when you picking up with replacement and without replace and will change if you draw two balls consecutively and without replacement then what is the probability that the first is blue and the second is white. So, but it will change here in this case, because the number of such blues which you can pick up in the first picking its 4 total number of balls is 10. So obviously, you can find out the corresponding relative frequency in the probability, but there is a slight change when you do it with replacement when you are picking on the second as white.

So, here this is the first one is basically blue so; obviously, white has not been affected. So, the second travel picking of which you do would have two whites, but the interesting part is that the total number of balls from where you are picking up the white is now 9 and not 10 because you have picked up one and basically removed it; that means, without replacement.

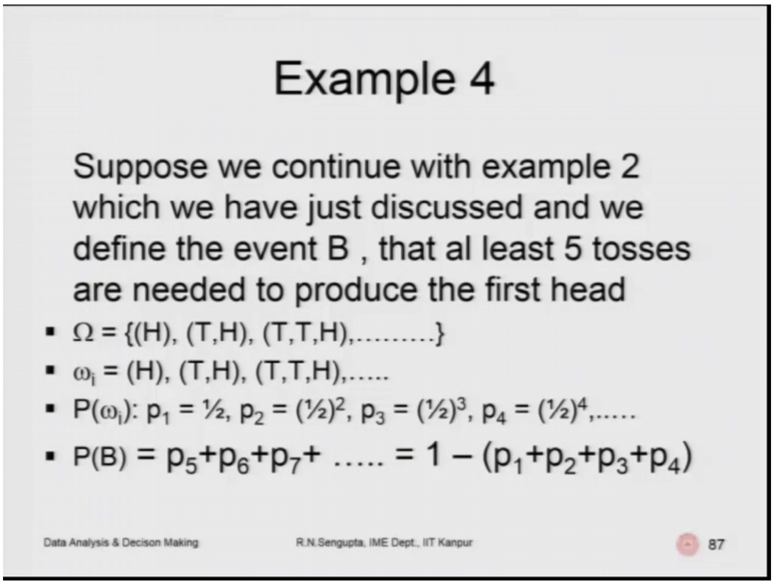
So, this concept should be clear. So, if I say that say for example, if you are picking up 3 balls one at a time with replace without replacement and you want to find out that; what is the probability that you will have all 3 whites; so the first one is white; that means, you have two balls out of the 10, you pick up that white and you remove it, then there is one more out of 9 you remove it, but the last case you do not have white; Obviously, but the total number of balls has now reduced by two which is basically 8.

(Refer Slide Time: 21:47)



We will consider the axiomatic definition of probability under this definition we consider the following the sample space is in infinite sample points are not equally likely which means that of the total number of sample points which basically makes the total sample space are infinite number. And obviously, with the corresponding probabilities of each sample points need not be infinitely finite.

(Refer Slide Time: 22:11)

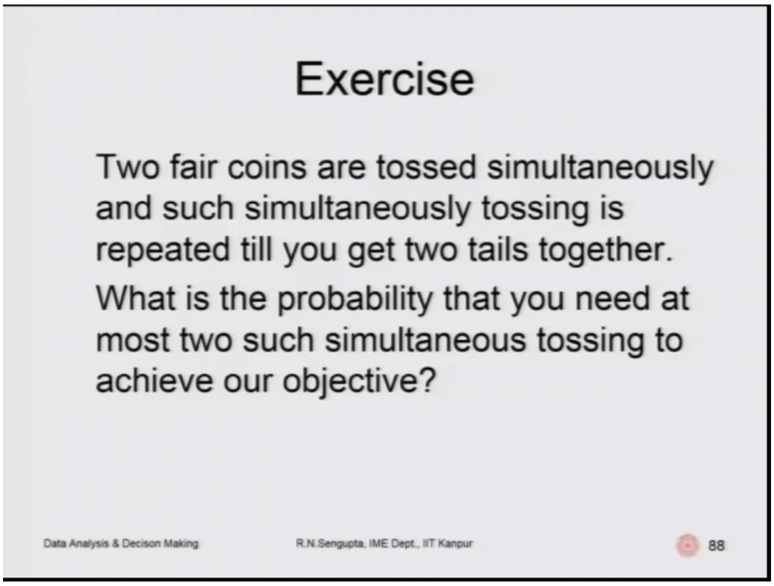


So, suppose that we are going back to the example which we did suppose we continue with the example two which we have just discussed and we define the event B that at least 5 tosses are needed to produce the first head. So, in that case what we have is that the overall sample space is now it is basically head, if you get a head, we stop if you do not get the head in the first row, we will basically have a tail and head. If we do not get a head in the first two it will be tail, tail, head, if you do not get head in the first 3, it will be tail, tail, tail, head and it will continue accordingly.

Now, in case when in that is the corresponding sample points would be given by as I mentioned is a head then the second sample point is tail head third sample point is tail, tail, head and so on and so forth. Now the corresponding probabilities when you of the sample points is that if it is the head in the first and nothing any stop experiment is basically half, if you get a tail and a head the corresponding probability is half into half which is half to the power square in the case, when you have to do with the head in the third picking it will be tail, tail, head; so, the corresponding probability half to the power 3 and so on and so forth.

So, if I want to basically have at least 5 tosses are needed to produce the first head. So, at least the 5 tosses means that you will have either 5 or 6 or 7 or 8 and it will continue till infinity. So, the overall probability would be 1 minus p 1 minus means you will basically first add up p 1 means that you get the head in the first p 2 means you get tail head p 3 means you get tail, tail, head p p 4 basically means that you get 3 tails and a head. So obviously, we will add up all this probability p 1, p 2, p 3, p 4, add it up. And basically subtract from one and get the corresponding probability which is for B two fair coin.

(Refer Slide Time: 24:02)

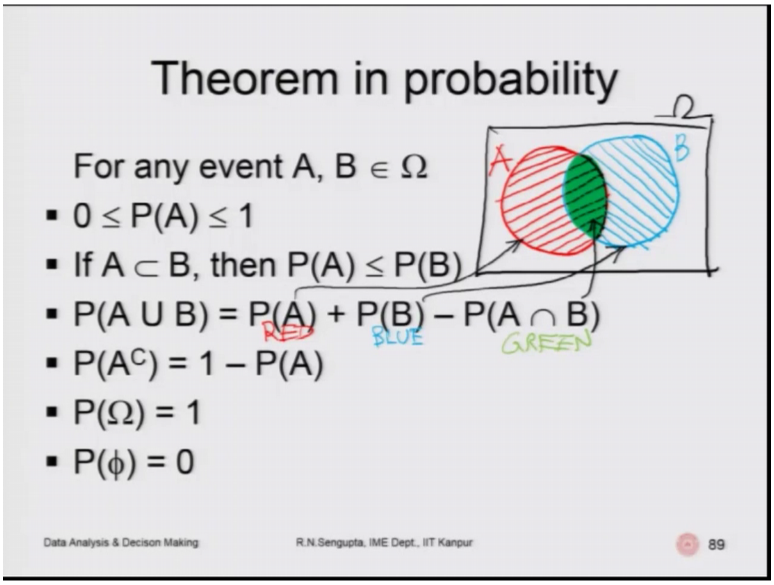


So, consider this two fair coins toss simultaneously and such simultaneous tossing is repeated till you get two tails what is the probability that you need to need at most two such simultaneous tossing to achieve our objective.

So, if you are basically scientist tossing and you want to get two tails together. So obviously, it would mean that head, tail or tail, head and head, head are not the combination. So obviously, only the combinations are tail, tail. So, in that case the corresponding probabilities would be if you get a tail, tail which probability half into half and the corresponding complimentary set would basically be head, tail, tail, head and head, head with basically is also equal to half to the power half and to half into half.

So obviously, in that case half into half means there basically; means half for tail head another half into half would be for head tail another half into half would be for the head, head and the last half into half would be for the case when you have tail, tail; so, if you continue doing it in this way, we can find out the corresponding probabilities and solve this problem.

(Refer Slide Time: 25:04)



Now, theorems in probabilities for any event A and B which is there in the overall sample space the overall corresponding probability of A is always between 0 and 1 inclusive. That means, the probability of A can be 0. So, probability, we can be one also if A is subset or a proper subset of B then the probability of A is always less than equal to probability B because probability is always increasing net it cannot decrease it can be 0.

But it cannot decrease. So, the universal set or the find trying to find out the of the universal of two events which is A and B is given by probability a plus probability B minus of probability a intersection B. So, what we are doing is this. So, this is A, this is B and apologize sorry for that. So, in this case let me use the black one. So, this P A is all the red one. So, it would be in red here P B the blue one and if I consider another color say for example, green for ease of understanding let me give the word of green here.

We use a bold color which is coming out to yellow yes. So, green this is the. So, you will understand that you want to find out the probability of A intersection B you have to find out the probability of a find out the probability of B and the intersection would basically B subtracted.

So obviously, if the intersection is null set it will be only be addition of that the probability of an a A compliment considering the event has been divided into two parts of the universal set is A. And a compliment would basically probability a compliment would be 1 minus probability of A, the third path of the second last point is; obviously, the probability of the overall universal set is one, because if this is the universal set. So obviously, overall probability here I am not marking it for the overall box is 1 and probability of the null set which is where there is no element is 0.

So, with this I will end the fifth lecture and continue more discussion about probability in the 6 and the corresponding lecture.

Thank you very much and have a nice day.