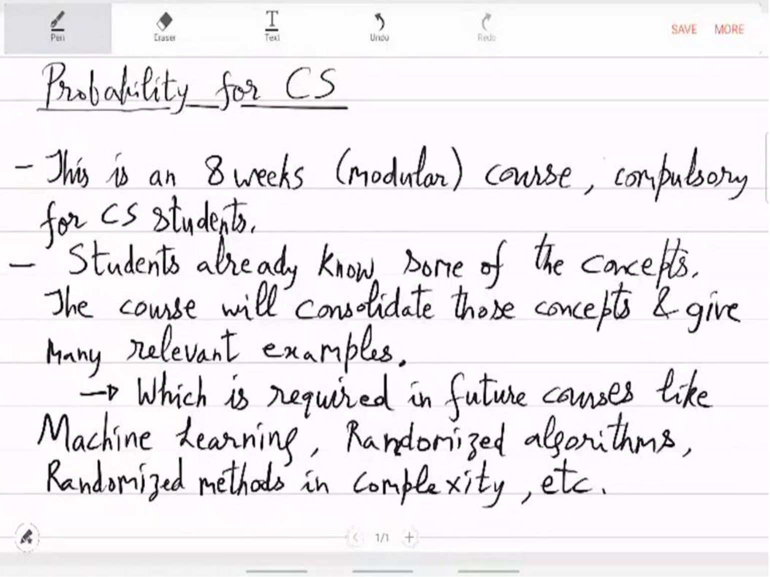
Probability for Computer Science
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology - Kanpur

Module - 1
Lecture - 1
Introductory Examples

(Refer Slide Time: 00:17)

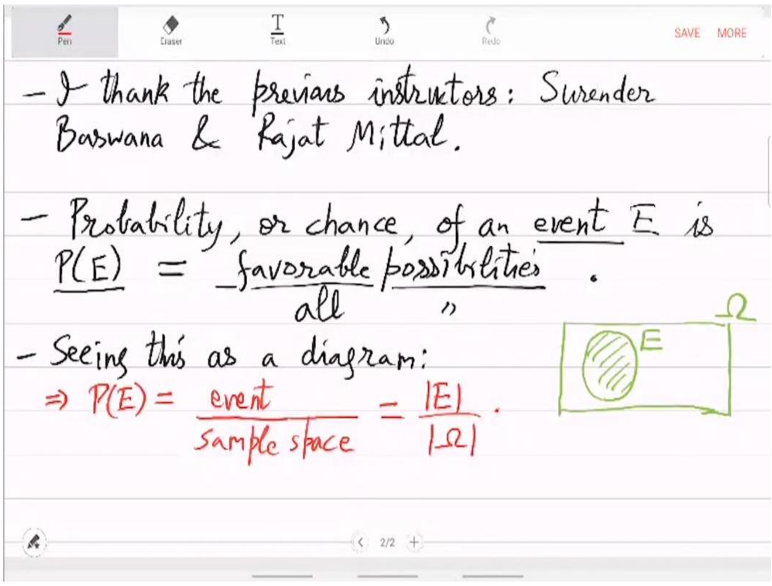


Welcome to this course on Probability for Computer Science. So, this is an 8 weeks long course only, which we call in IIT modular course. So, instead of 12 weeks, this will be for 8 weeks, and this is compulsory for CS students. So, students, specially our students who have given JE, they have already learnt probability and how to apply it in fairly complex problems. So, you already know some of their concepts.

So, why do you do this course again? Well, because this will consolidate those concepts and this will give you far more examples to think about and to remember and then to apply. So, the course will consolidate those concepts and give many examples. So, these examples will obviously require understanding of math, but they will not be purely mathematical. They will be inspired from real life situations that appear in computer science, problems inspired from.

Also, it is a compulsory course, because it is useful in, it is in fact required for more advanced courses, future courses like machine learning, randomised algorithms of course. So, randomised algorithms, obviously they use probability both to run the algorithm and then to analyse it's complexity and success or failure, how good is it. And randomised methods in complexity. So, whether you are interested in theory or in practice, in both kind of courses probability will be omnipresent.

(Refer Slide Time: 03:39)



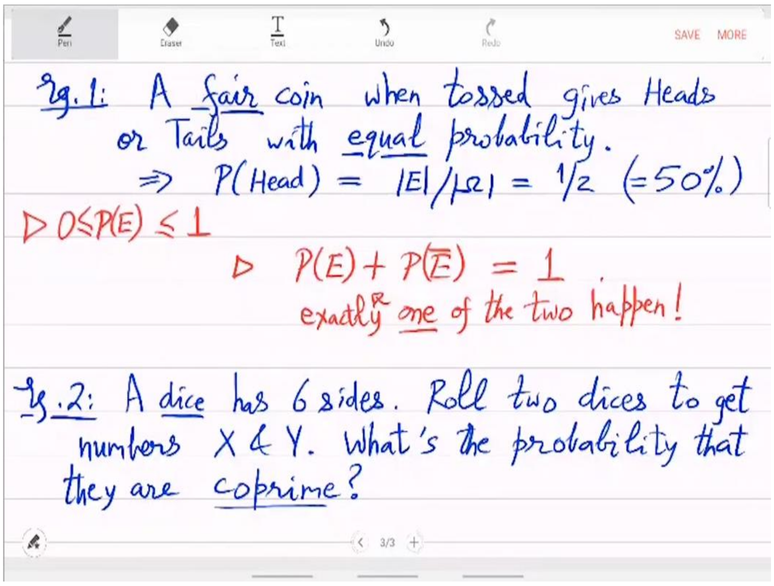
I would start by thanking the previous instructors who have taught this course. So, Professor Surender Baswana and Professor Rajat Mittal. So, much of the material will be based on the standard syllabus of this course. So, just to start off, what is probability? or what is chance? So, probability or what we simply call chance of an event is; so, for now, we will not define it very properly, what is an event? what is a probability? but you, since you already have some ideas, so, let us just continue with this intuitive notation.

So, probability of an event E which is P of E is essentially number of favourable possibilities divided by all possibilities. So, it is the ratio of favourable over all. So, if you see this as a diagram, Venn diagram to be proper, then it is just; so, there is this universe or the whole sample space which we call omega. And in this, there are these favourable possibilities, which is the event inside the circle.

So, you are interested in this. So, how big is this space in comparison to the whole sample space omega? So, this ratio is probability. So, if things were, if your experiment was completely random, there was no bias, whatsoever, then, whatever this Venn diagram says mathematically is what you will expect intuitively to happen in practice. So, theory and practice will meet.

If everything was ideal, then this probability is correct, which is size of E over size of omega. So, it is event over the sample space, which in numbers is size of E over size of omega. This is what you know from your school days. There is nothing new in this. No surprises. So, let us start with some examples which will become more and more interesting as we proceed.

(Refer Slide Time: 07:22)



So, what is the most basic example? In fact, the example which must have started probability, which comes from gambling. You go to a casino and you play a game which is a game of chance. So, somebody flips a coin. And based on what appears in the coin, either you win or you lose. So, that is really the starting point of probability itself historically. So, it starts with a fair coin.

So, a fair coin when tossed gives heads or tails with equal probability, which means that; since there are only 2 possibilities and say you want heads. So, if you want the probability of heads in the experiment where you have flipped a coin, single coin, fair coin, it is also called unbiased coin. So, that will be E over omega, which is 1 over 2. So, this, you can also write as 50% chance.

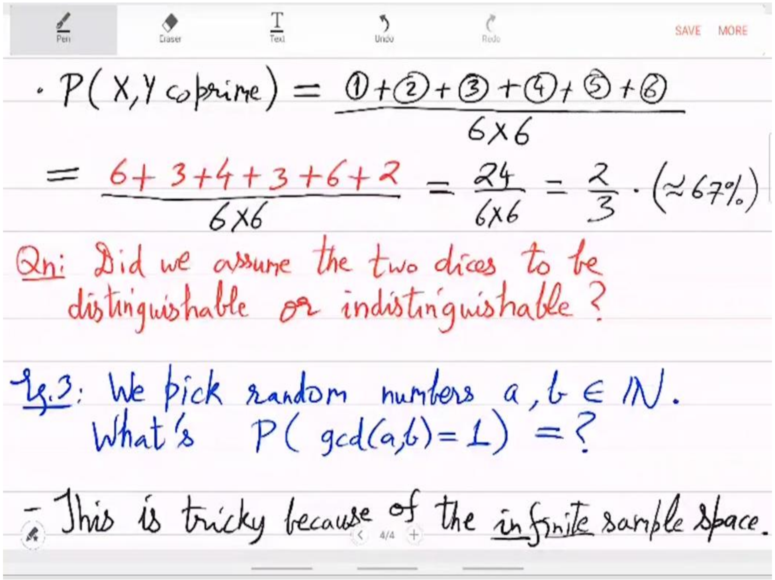
But we will always use fractions, so, we will call this half or 0.5 probability. But sometimes we just multiply it by 100 and we call it 50%. So, this is coming from equal probability. That is the definition of unbiased. What you remember already is that probability is at most 1, because it is a ratio of favourable over all. So, it cannot be more than 1. And it is always non-negative. It cannot be negative, because you are actually counting.

So, that is the first property. And second property, from the, again, look at this diagram, Venn diagram. From this, you can see that E and its complement fill the whole space, which we write as probability of E and probability of E complement. That sums up to 1, which is essentially saying that one of the two things will happen. In an experiment, exactly one of the two happen. That is the way you can interpret physically this equation.

So, these 2 things are both intuitive and easy to show by whatever definition you use of probability. Let us make it slightly more interesting and let me say that a dice has 6 sides for me; so, numbered 1 to 6. So, roll 2 dices to get numbers X and Y. Now, what is the probability that these 2 numbers are coprime? Coprime means, they do not share a factor other than 1.

So, for example, 1 and 2 are coprime but 2 and 4 are not coprime, because 2 and 4 share two. So, you get 2 numbers between 1 to 6. What is the chance that they are coprime? So, how do you solve this? Given your basic understanding, you count all the possibilities.

(Refer Slide Time: 12:04)



So, probability that X, Y coprime. So, you look at the case. Let us divide this into X = 1, X = 2, X = 3, X = 4, X = 5 and X = 6. So, there are 6 values X can take. And then we will later on count how many Y's are possible. That will give us the favourable possibilities or favourable outcomes. What are the total outcomes? So, total is X up to 6, Y up to 6. So, it is 6 times 6. So, now, if X is 1, then Y can take any value. So, there are 6 values.

If X is 2, then Y can take 3 values. Then if X is 3, it is only 3 and 6 which are bad. Everything else, Y can take. So, it is 4 possibilities for Y. If X is 4, it is like 2, which is again; so, Y cannot be 2, 4, 6. So, it is 3 possibilities like before. And if X is 5, Y can be anything. And finally, if X is 6, then Y can be only 1 or 5. So, that is 2. So, this divided by 6 times 6; which is what? So, that gives you 24, 6 times 6, which gives you 2 thirds, which is around 67%.

So, when you roll 2 dices, the chance is actually very high that the numbers will be coprime. That, say that which is something you would not have guessed immediately. You get that only after doing this rigorous calculation. So, you can see that probability in these examples is just counting. You have to count carefully and then you get a sense of how probable the event is. So, let me ask a question here in this calculation.

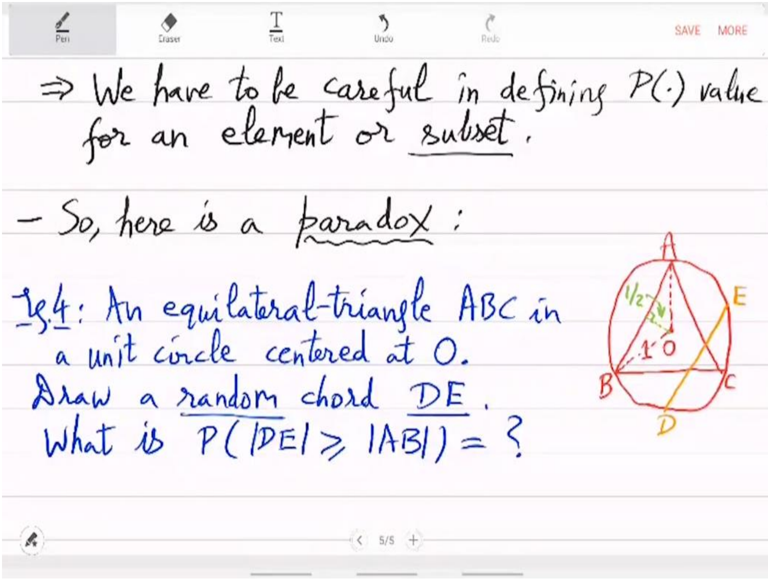
So, in this calculation, did we assume the 2 dices to be distinguishable or indistinguishable? What do you think? When we wrote the favourable possibilities, what did we assume about the dices? Are these 2 dices the same or are they different? So, think about that, accordingly; so, in one case you will get one answer, in the other case you will get a different answer. Work this out as a homework exercise.

Now, building on this, we can ask the question; comes our third example, that suppose from this infinite expanse of numbers, we pick 2 numbers at random, what is the chance that they are coprime? So, suppose you pick random numbers a and b; they are positive integers, so, 1, 2, 3... That is what this N denotes. What is the chance that they are coprime? So, what is the probability that the gcd of a, b is equal to 1?

So, this is a question similar to what you just saw, but difference is that, there a and b are between 1 and 6. So, you picked something out of the 36 possibilities and you did the explicit calculation. But what will you do here? Here now, it is not just 6 possibilities, it is infinitely many possibilities. So, does it make sense to even talk about probability? which forces you to formalise probability in a more general way.

So, it is not; somehow, at this point the intuition breaks down and math then has to take over. So, this is tricky because of the infinite space. So, if you look, if you just do 1 over infinity, that will be 0. Then everything, every event seems to be probability 0. So, can we make a more refined definition of probability than just ratios?

(Refer Slide Time: 18:03)



So, this means that more formalism is, more careful formalism is needed, which you may not have done in your high school. So, that is where this course will depart from your earlier knowledge. So, we have to be careful in defining P, the value of probability for an element or subset. So, think of every event as a subset of some space. But when that space is infinite, then you have to be more careful about how to define P.

So, here is a paradox. We will see some problems that happen with infinity. So, that becomes our example 4. So, let us draw a picture first. So, that is a circle centred at O, say it is a unit circle, equilateral triangle ABC. So, if you look at the radius; since this is 1, and since the radius is 1, the angles are 60 degree. So, you can use sine and you can deduce that. This distance AB from the centre O will be half.

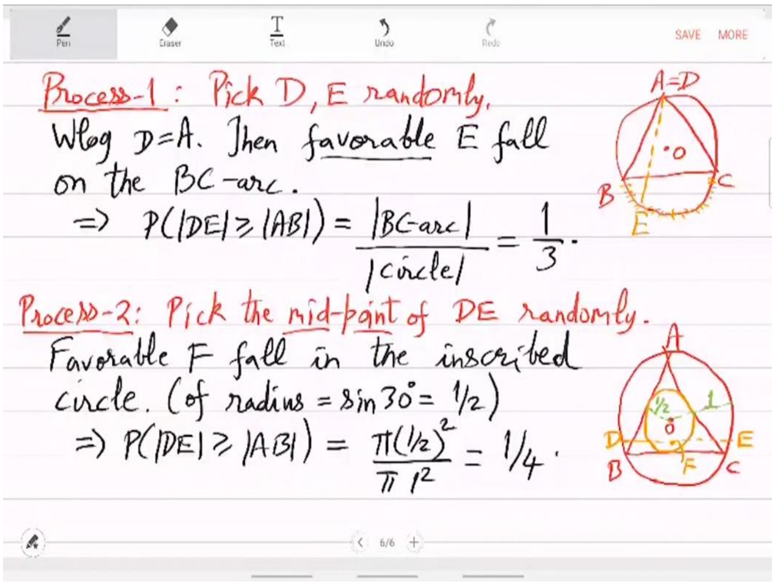
And you can also deduce what is the length of AB. All that is pretty standard. So, what is the question? The question is that, given this picture, if you draw a random chord, So, let me draw it. That is a random chord; two endpoints D and E; what is the chance that this chord has length more than AB? What is the chance that this, or random chord is longer than the edge of the equilateral triangle? This is what you have to calculate.

Now, the number of possibilities here is mind boggling, because it is, we are really talking about the real picture. So, infinitely many D's, infinitely many E's, infinitely many lengths.

So, the traditional definition of ratio will just give you something over infinity, all possibilities infinity. So, this is like always 0, you would think. But we can still make sense of this as follows. We will actually compute probabilities now, of this event.

An equilateral triangle ABC in a unit circle centred at O. Draw a random chord. We called it DE. So, what is the probability that the length of the random chord is more than the length of AB, let us say at least? What is this probability? Does it make sense first of all to ask for this probability? And then second, how will we calculate this thing, where the possibilities are endless, literally infinite? We will make sense of it not in one way, but actually 3 ways. You will see it now.

(Refer Slide Time: 22:58)



So, let me define process 1. Process means how are you selecting randomly the chord? There are many ways. So, you can pick D and E in a random fashion. Let us do that. Now, since this is a symmetric picture and you are selecting 2 random points; so, without loss of generality, you can assume that D and E are the same. And then, you only have to select E randomly. So, with this picture in mind, what are the good E's?

The good E's are the ones which are in; so, for example, if you take E in the AB arc, then the chord will be smaller than the edge. And the symmetrically AC arc will be bad. So, the only good arc is BC. So, these are all good, favourable possibilities. So, let us write that down. So, pick D, E randomly. And to work this probability out, it should not matter if you fix D, because the picture is symmetric.

So, without loss of generality, D we can take to be A. Then favourable E fall on the BC arc. So, then you just have to compare the circumference. Which means that the probability of DE being at least AB, this is equal to the circumference of BC over the circle, which is clearly one third. So, we have a probability. If you do, if you follow this process of picking D and E randomly, then without going into the formalism of probability, it seems that one third is the correct answer, because selecting DC arc should correspond to its length.

And length is not infinite. Length is just, you can get it by 2 pi formula. So, length is just 2 pi by 3. So, from that, you get this finite number. But this is not where the story ends. So, you could have chosen the chord in a different way. So, for example, what you could have done, pick the midpoint of D. Note that if you pick the midpoint of D, then you all, then it specifies D and E uniquely; it specifies the chord, because the end points are defined by the intersection with the circle.

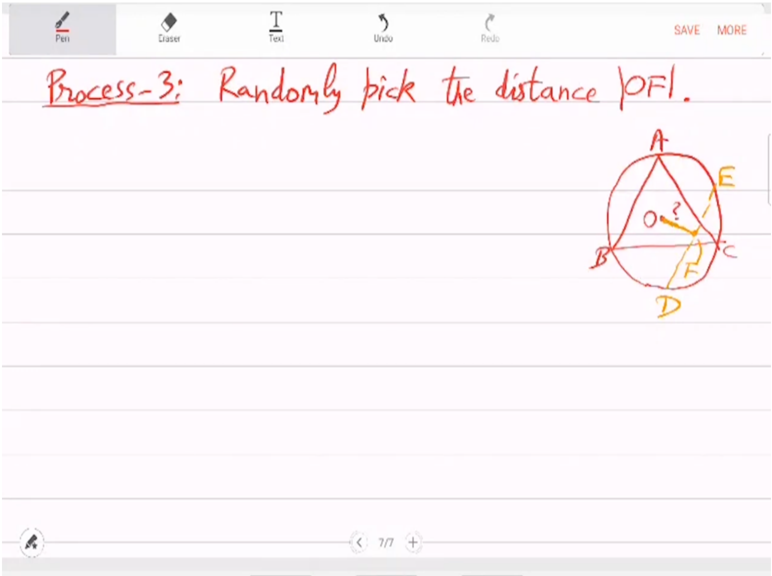
So, picking the midpoint of chord is also a valid selection process. So, then what happens? What are the favourable possibilities from that point of view? In this case, I leave it as an exercise that this inscribed circle. So, what is the radius of this circle? So, the radius is half. So, the inscribed circle is radius half, circumscribed is 1. So, if you pick a point here, for example this point; so, then you draw the chord. That is your DE. And let us call this point F.

So, by first picking a random point F and then picking that chord whose midpoint F is, the chord will be longer than AB if and only if F falls in this inscribed circle, the favourable F fall in the inscribed circle. So, this you can prove geometrically. It is a simple exercise to do. I will not do it rigorously here. But it looks believable. If you look at the picture, then F which is outside the inscribed circle, those chords will be smaller.

And F inside the inscribed circle, since it is closer to the centre O, the chords will be longer. That is basically the visual insight, visual proof. And by that, so, inscribed circle whose radius is half. This half came from sine 30 degrees, which means that the probability that the chord is at least the edge AB. This comes out to be the areas of the circles. So, F, favourable is in the area pi r square and total area is pi. So, which gives you 1 by 4.

So, now you are getting a, this gives you a different probability. This now comes out to be smaller than before. Although each of these processes, they seem valid, they seem a good way to pick a random chord. But then, the way you are picking it, you are getting different probabilities. And even this is not the end.

(Refer Slide Time: 30:48)



We could do it in a third way, which is, pick the distance of DE from the origin. So, randomly pick the distance which will be OF. So, that means you will, this distance of O from F, this is what you are picking in a random way. So, not the point F, but the distance which is OF. So, it seems that it is the same as picking F, but you will see that probability will come out to be different.