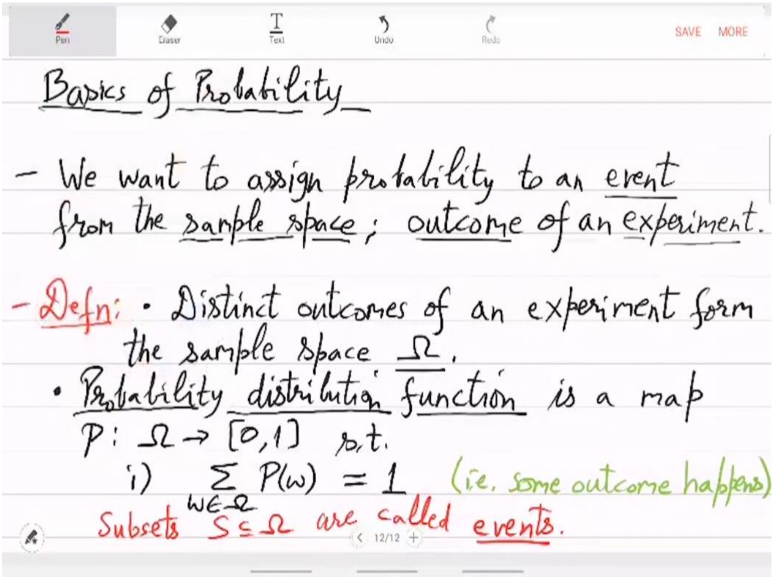
# Probability for Computer Science
**Prof. Nitin Saxena**
**Department of Computer Science and Engineering**
**Indian Institute of Technology - Kanpur**

## Module - 1
## Lecture - 3
## Probability over Discrete Space

Last time we gave the outline and now we will start with basic concepts of probability.

**(Refer Slide Time: 00:19)**



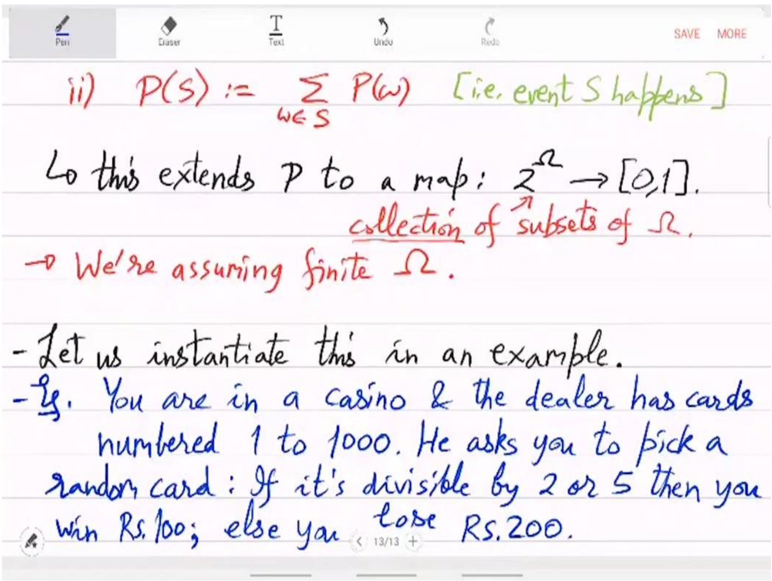

So, what do we want to do? and then, what do we want to formalise? So, we want to assign probability or this quantitative notion of chance to an event from the sample space. In other words, many events are possible, many events may happen, and we want, we are interested in some event and we are there interested in asking or quantifying what is the probability of that particular event happening.

So, these are the 2 key words, event and sample space, but you can also talk about outcomes of an experiment. So, either event from the sample space or outcome of an experiment, you want to assign probability. So, let us give first definition and see some interesting examples, before we make even more formal. So, let us say that distinct outcomes of an experiment, they live in a sample space omega.

Or, if you collect all the outcomes possible, then they form the sample space omega. That is the universe in a problem. So, once you understand the universe and elements in it, the next is probability distribution function. So, what is that? So, that is a; so, this long thing, we can just call it probability or probability function. So, probability is a map P from omega to 0, 1. So, every element in the sample space, every outcome is assigned a real value between 0 and 1, such that obviously there has to be some interesting axioms, otherwise this would not be, probability would not be interesting enough to study.

If it was just a map, that is not interesting. So, you have to put some conditions. So, what are the conditions? First you want that the probability summed up for every element in the space, that should be 1. So, this, remember this captures the idea that something has to happen. So, that is, some outcome happens. So, that idea is captured by saying that sum of probability is 1. It cannot be that nothing happens. That is what this is capturing. So, subsets S of omega are called events.

**(Refer Slide Time: 04:45)**

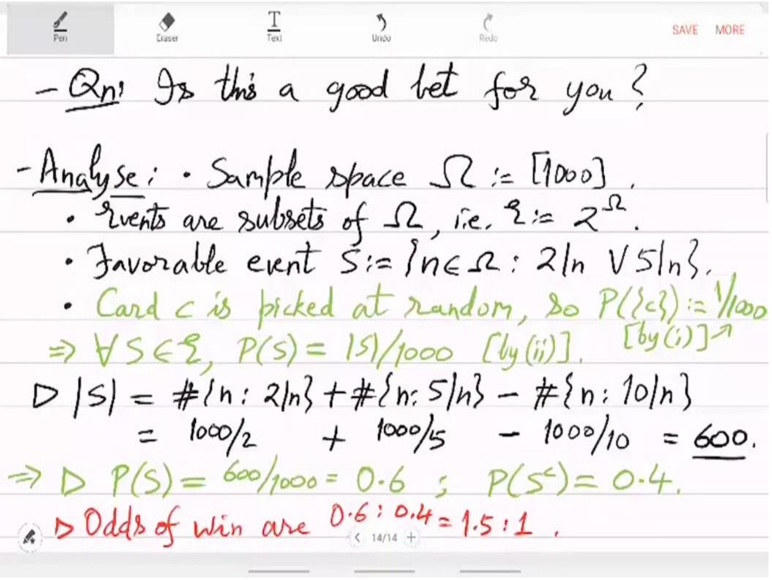



And what we want for these is probability to be extended. So, for every subset, any subset S, P of S is defined as sum over the probabilities of elements omega in S. So, this is like saying, that is S happens. And note here that this extends P to a map from subsets of omega 2, 0, 1. So, this is the probability function. It not only assigns 0, 1 values, 0 to 1 values to elements, but also to subsets. So, these are the subsets of omega. So, collection of subsets.

So, probability is a map from that to 0, 1 on the real line. This is the full definition. And since we are doing a sum, so we assume finite omega. Otherwise, if S is infinite in point number 2 or if omega is infinite in point number 1, then the sum, how do you do the sum? Infinite sum is not really defined. Then you get into ideas of, questions of convergence and divergence. So, we will postpone that. For now, assume that omega is a finite sample space.

So, we give probability definition via this, these 2 sums. So, let us instantiate this in an example. So, suppose you are in a casino and you want to play a game, where the dealer who has cards, they are numbered 1 to 1,000. So, what is the game? So, he tells you to pick a random card. And if that card number is divisible by 2 or 5, then you win 100 rupees. Else you lose 200 rupees. This is the game in a casino. Now, you have to decide whether to play this game or not? Is it a good idea to play this game?

**(Refer Slide Time: 08:44)**




So, that is the question. Is this a good bet for you? So, for this decision making, you have to use probability. You have to see what are the odds of you winning? If the odds are high, essentially, here the, here you are, if you win, you only get 100, but if you lose, then you lose double. So, intuitively, your decision will be to play this game only when the odds are more than double of winning. So, is that the case? So, let us check that.

So, let us do a probability analysis. So, formally, what is the sample space? So, sample space omega is the cards that are possible, that you can pick, which is 1 to 1,000. So, this set, 1 to 1,000 is the sample space. What are the events or outcomes of the game? So, events are subsets of omega. That is, E is 2 raised to omega. And what event are you interested in? So, favourable event is that subset of numbers where either 2 divides n or 5 divides n.

That is the favourable event. What is the probability distribution now? So, card C is picked at random. So, the probability of this singleton set C is 1 in 1,000. This is because the sum of the probability should come out to be 1. Now, there is, since there are 1,000 elements, so you get the probability to be 1 over 1,000. So, that is by property 1. That was the property that P has to satisfy, sigma P = 1. So, from that, you get 1 over 1,000.

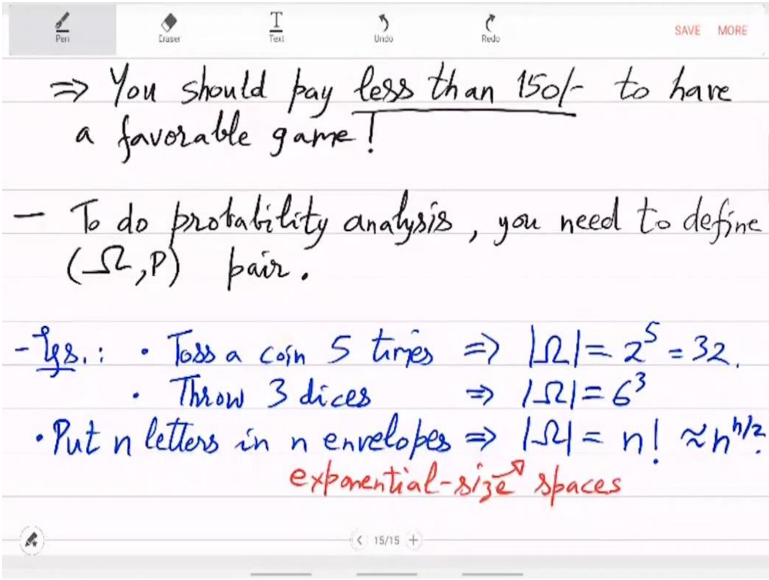
And so, this means that for every subset in E, probability of that event S is, again you would sum up, so you will get size of S by 1,000. This is happening by property 2. So, that is the instantiation of the notation that we have defined. So, what is omega? what is big omega? what is E? and what is the favourable event S? and finally, what is the probability map on these? But we still have to do the calculation to find odds of winning in this casino game.

So, for that, we have to calculate the size of S, which is number of n's divisible by 2 or 5. So, we can break this up into multiples of 2. So, number of n, such that 2 divides n. Then number of n such that 5 divides n. But then, numbers that are multiples of 10, they are being double counted. So, let us subtract that. So, numbers, n such they are, which are divisible by 10. So, how many numbers are there divisible by 2, between 1 to 1,000? That is 500.

Those divisible by 5 is 200. And those divisible by 10 is 100. So, that gives you the number 600. So, that is your count. 600 numbers out of 1,000 are either divisible by 2 or by 5. So, it seems that they are more than 50%. Does that mean you should play the game? Not really. You have to look at 600 versus 400. So, what you get is, probability of S is 600 by 1,000, which is 0.6 and probability of S compliment is 0.4.

So, odds of win are, 0.6 is 2.4, which is 1.5 is to 1. So, odds are 1.5 is to 1. It is not 2 is to 1. So, you should never agree to pay or lose anything more than 150. Anything more than 150 rupees would mean that it is a bad game for you. So, the casino is tricking you. So, usually, in this game, people will be paying the casino.

**(Refer Slide Time: 15:50)**




So, your decision will be, do not play the game. So, this means that you should pay less than 150 rupees to have a favourable game. And since the dealer is asking you to pay 200 on losing, you decide not to play this game. So, this was not clear in the beginning, but when you do the analysis, probability analysis, it becomes very clear. So, hopefully this definition and application is clear. So, let us now move to few more points.

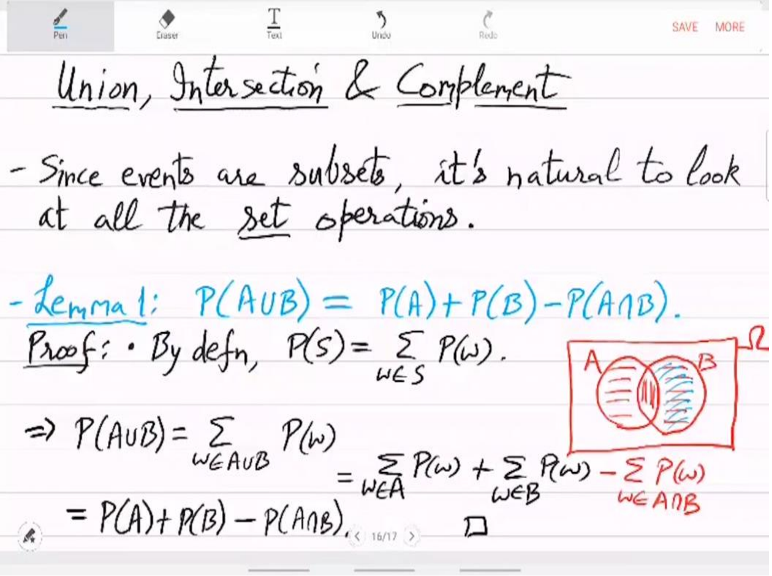
So, you need to define. So, to do probability analysis, you need to define omega, P pair. So, what is the sample space and what is the probability map on the subsets, on the subset space, satisfying those natural conditions that you want from probability. And we have covered the definition for finite omega. Infinite omega, we will see later. That is a bit more abstract. So, more examples. So, if suppose you toss a coin 5 times. So, what is omega?

In particular, what is the size of the sample space? What are the possibilities? So, that, one time you get 2 possibilities, heads or tails. So, it is two 5 times. So, it is 2 raised to 5. Suppose you throw 3 dices, so, what is the size of the sample space? 1 dice has 6 possibilities. 3 times means 6 cube. So, you get 6 cube. And finally, put n letters in n envelopes. Now, what are the possibilities here? So, 1 letter will go in its own envelope. There are n envelopes.

But then, if you are not careful, you will make a mistake. So, you, or letter may go in any of the envelopes, but 2 letters will never go in 1 envelope. So, it is basically a bijection. So, how many bijections are there? That is the same as number of permutations on n letters or n symbols. So, you get n factorial. And n factorial is very big. It is something like n to the n by 2. So, if you have 100 envelopes, then the sample space is huge.

It is an exponential sample space. The same thing happens with previous examples also. If you toss a coin n many times, then you get 2 raised to n possibilities. And if you throw a dice n times, then again you get 6 to the n. So, these are all exponential spaces. So, usually the sample space is very large. And in that large sample space, you are looking for favourable elements, and so, favourable events. So, with this, let us move forward to what the probability properties imply, what can you do with it, operators on probability.

**(Refer Slide Time: 20:43)**




So, let us study union, intersection and compliment of events. So, since events are subsets, it is natural to look at all the set operations. So, as the set changes, how will the probability change? We want to understand that, the action of probability on modifying the subsets or events. So, towards that, let us prove an interesting property which will be very useful. You already use it all the time, but let us see it formally.

So, probability of a union, what is that? A and B are events, so, what is probability of union? Now, you would expect, it is to be sum of the 2 probabilities, either A happens or B happens. So, you may write as a first draft P A + P B, but this is not quite correct, because maybe there are possibilities, elements that are shared between A and B. They are been double counted. So, you subtract that. Then, this looks correct.

So, let us formally prove this, which essentially is this picture. So, this is the sample space, this is your A, this is your B. So, what is happening is, this intersection. The intersection is being double counted. So, the shaded part; basically, there are 3 shaded parts, so, you want to compute the probability of that event. So, intuitively it is clear, but we have to formally prove it using the definition of probability distribution function.

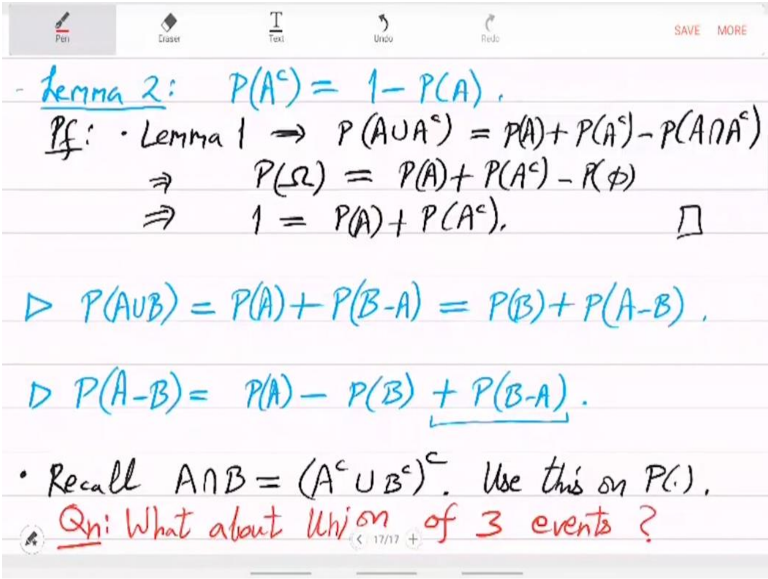
Remember what the definition was. So, the definition was simply by these 2 sums. So, first was sigma P = 1, over a all elements. And second was that probability on subset is sum of probability over the elements. So, you just have those 2 things? So, using those 2 things, what can you do? How do you prove lemma 1? There is no other physical interpretation of probability allowed. Let us write the definition.

So, by definition, you know that probability S is sigma P omega elements of S. So, which means that probability A union B by definition is omega in A union B sum of P over all the elements A union B. Now, this sum can be broken into 2 parts, omega in A, omega in B. But then, this is not quite correct. The sum we have written is not correct, because there are these omegas which are in both A and B.

So, you are double, you are adding them in both the summons, in both the sums. So, we have to remove one of those. And this is where subtraction happens. And with this, again, using the definition, you get that the first sum is probability of A, second is probability of B and third is probability of A intersection B. So, it is a proof that is completely unsurprising, but this proof still has a point which is, it does not use any physical interpretation, it is really, you can write this proof just from the definition of the map P.

Just from P of S equal to sigma P, you can prove the lemma. So, we have formally proven that probability distribution function always satisfies this property, when the sample space is finite. Now, this leads to numerous other identities, connections.

**(Refer Slide Time: 26:50)**




So, first is very simple. It is about the compliment. So, probability of A compliment is 1 minus probability of A. Again, you are not allowed to use interpretation of probability. You have to prove this syntactically from, directly from the definition of P map. But that is fine. What we can now look at is A union A complement, and use lemma 1. So, lemma 1 implies that probability of A union A complement is probability of A plus probability of A complement minus probability of A intersection A complement.

Of course, A union A complement is everything. So, you get probability of omega equal to probability of A plus probability of A complement. What is A intersection A complement? That is empty. So, clearly, probability of empty is 0. This we could have actually put in the definition. Probability of empty has to be 0. And it is actually already in the definition, because probability of a subset was defined via sigma P.

So, maybe we could go back and write that a bit more clearly. Here we can write. So, probability of empty is 0. From this, we have all the values now. So, this means that 1 is equal to probability of A plus probability of A complement. So, both properties 1 and 2 are invoked in lemma 2, which invokes lemma 1. So, ultimately both properties 1 and 2 are used to prove this fact, which is completely consistent without physical interpretation, but that is not needed.

This is our rigorous proof. And one more thing I can write, which is that probability of A union B is what? Probability of A plus; so, what is this probability of B minus probability of A intersection B? what does this denote in the diagram? In this Venn diagram, this is from B you are removing A intersection B. So, it is actually this part. So, this is, in other words, B minus A. And since A union B is symmetric, you get also probability of B plus probability of A minus B.

So, we also have a formula now, another formula we have that probability of A minus B is equal to probability of A minus probability of B plus probability of B minus A. So, this is the part which has been added. So, probability of A minus B is not just difference of the probabilities. Obviously, that will not be true because then it is possible that right hand side is negative, in case probability of B is higher. So, correct formula actually adds this B minus A also.

So, we have now formulas for complement, union, difference and in a way, you can also get for intersection. So, recall that A intersection B is nothing but A complement union B complement, complement. So, using this, you can also give a formula for intersection. But generally, that is not very helpful. So, this is not very useful, but you can get something out of this.

So, those are the simple nice formulas they follow from the formal definition of P map. So, next, what we can ask is the question, what about union of 3? So, union of 2 events, you understood, what about union of 3? then, what about union of 100 events and so on. So, that, we will do next.