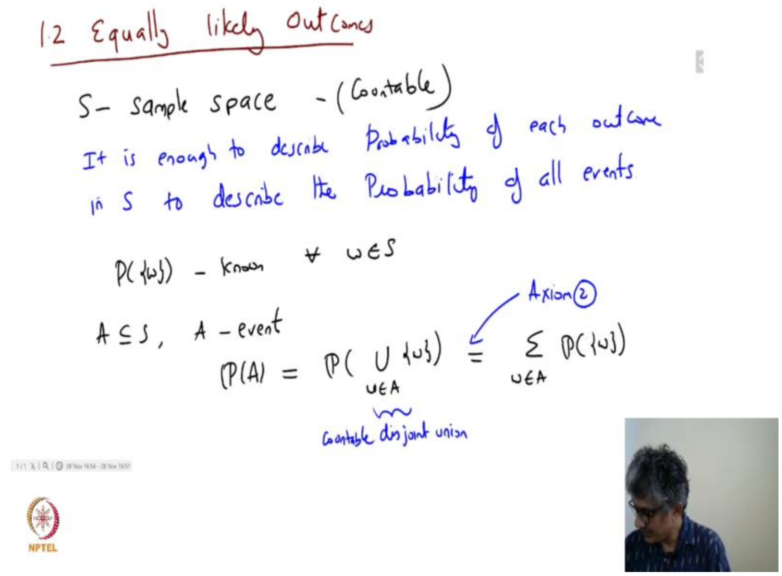
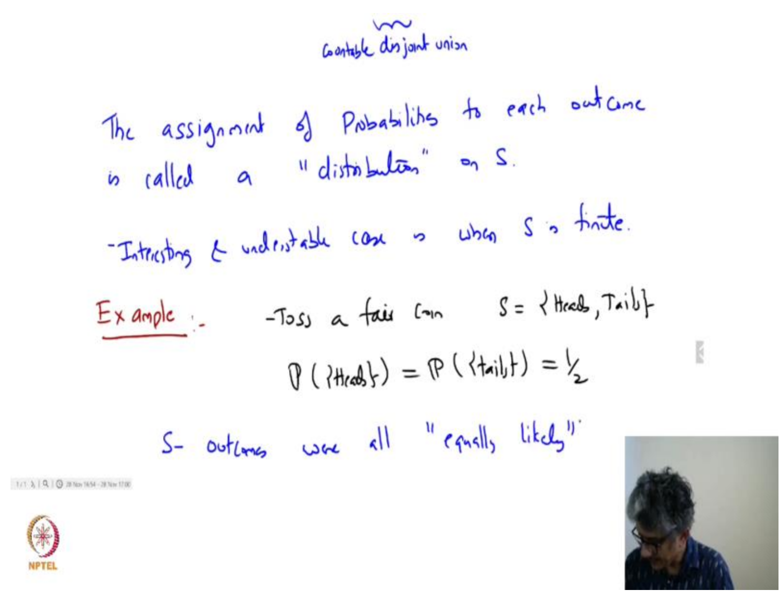
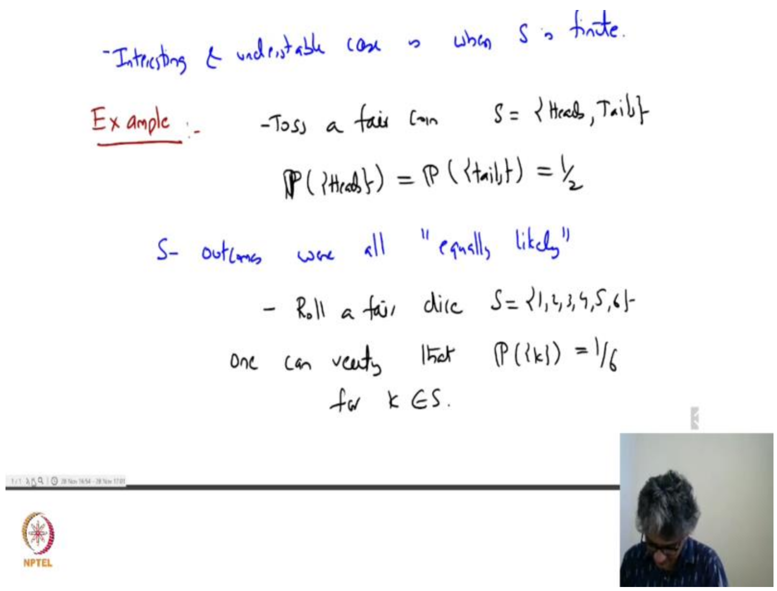
Introduction to Probability-With Examples Using R
Professor Siva Athreya
Theoretical Statistics and Mathematics Division
Indian Statistical Institute, Bangalore
Lecture – 03
Equally likely Outcomes

(Refer Slide Time: 00:18)







So, today we will do a very important set of sample spaces or a concept you can call it; I will title as that equally likely outcomes. So, here the idea is that you have a sample space; let us say it consists let us say S sample space. And S is countable that is you can count the elements, it says countable collection; you can think of a finite set or some index that are actually numbers. So, one important feature is it is enough to describe the probability of each outcome in S to describe the probability of all events. So, by this I mean the following that let suppose I know I know probability of omega, for I know this this is known for all omega in S.

Then if you look at any event A; A subset of S, A is an event. Then the probability of A, you can write it first as the probability of union of all omega in A of the sets omega. Now, this is either of the disjoint union of countable sets disjoint union, and I just write as countable disjointed union; because S is countable and A is subset of S. This is equal to now the sum over omega in A the probability of omega, by the Axiom 2 of them; because I am using Axiom 2. So, now you notice that moment I know probability of every single outcome; I know the probability of the event A.

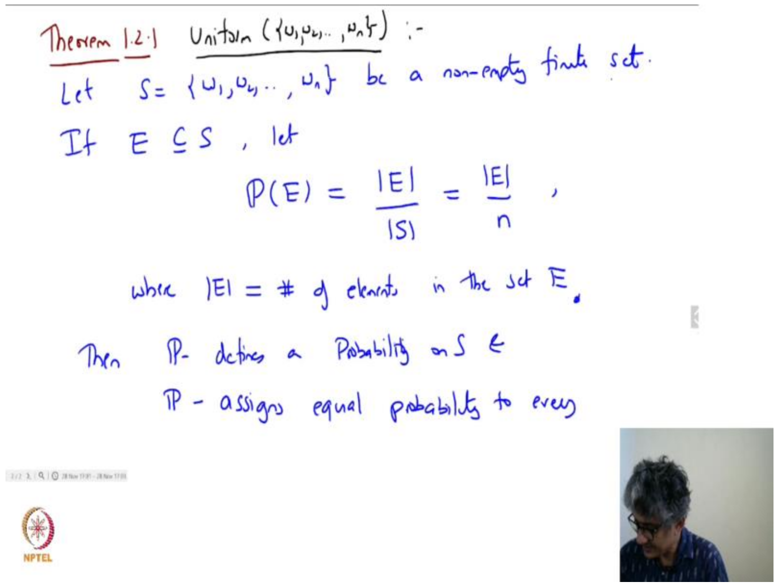
So, in some sense the assignments or probability of each outcome gives sort of a kind of a weights for every outcome; which will describe as the (prob) as the distribution. So, this is a special term that people use all the time. The assignment of probabilities to each outcome is called a distribution; a distribution on S is called distribution on S. So, that is this is distribution as means it describes how probabilities are assigned to each part. So, in the countable setting itself is this is generality it is true; but the most interesting reason interesting case occurs in interesting and understandable case is when S is finite.

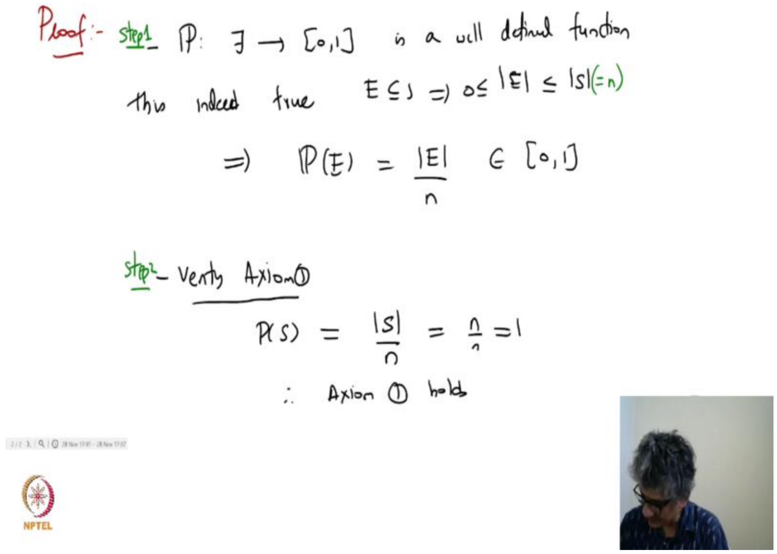
And here this is sort of a unique feature; so let us take two examples, I will come back to the examples again we draw (())(04:38). Let us take one example let us say as an example; so one is that let us say we toss a fair coin. And this is an experiment in which S is heads and tails; and we already saw that the probability of heads in this case was equal to probability of tails, was equal to half. So, here was a here was set of when S was finite; tossing of fair coin it means that heads and tails are equally likely to apply. And that gave me the probability of heads equal to half and tails equal to half.

So, just to rephrase this this sort of this example that we did before; so this is the case in which S the outcomes in S were all equally likely. That is every outcome is is half probabilities, if there are two outcomes each outcome (probability). This phenomenon is a little bit more general; for example let us go to the other example.

Let us say we roll a fair dice, so the outcome is just S will be 1, 2, 3, 4, 5 and 6; and then we can verify one can verify that the probability of any single outcome is actually one sixth is one sixth for all k in S. So, again an experiment where we see that the the probability of any single outcome is same as any other outcome in the sample space. So, this phenomenon is little general so in that sense; so we try and place this phenomenon in a general setup, when S is any constitute of n elements, so n is arbitrary. So, I will write in the form of a theorem.

(Refer Slide Time: 07:08)





Here is the theorem, so I will call it 1.2.1; and then distribution or the way I assign probabilities to the set is called uniform of omega 1, omega 2 up to omega n. So, that is the that is the that is what the theorem is going to do; so the idea is that you let S be finitely many elements omega 1, omega 2 up to omega n; n elements so this be a non-empty finite set. Then if E is a subset of S, let the probability of E to be equal to the number of elements in E, divided by the number of elements in S. The same as the number of elements in E divided by n. So, let us say where E where mod E is the number of elements in the set E; that is how I am going to define a probability.

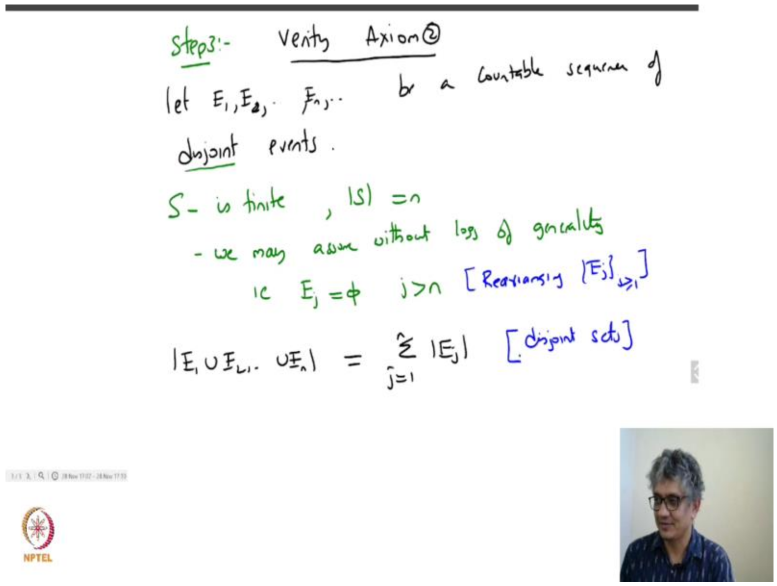
Then the conclusion is then so E is this that this vector is; then the conclusion is that let me full stop here. Then the conclusion is that P defines a probability of S probability of S, and more importantly the crucial thing is that P assigns equal probability to every individual outcome in S; P assigns equal probability to every individual outcome in S, so that is very crucial one. Let me do the proof of this; so what do I mean by this? The this theorem. Let us just before write the proof let me explain the theorem again. See I have a sample space S and what I do is I have the event space instead of all subsets of S.

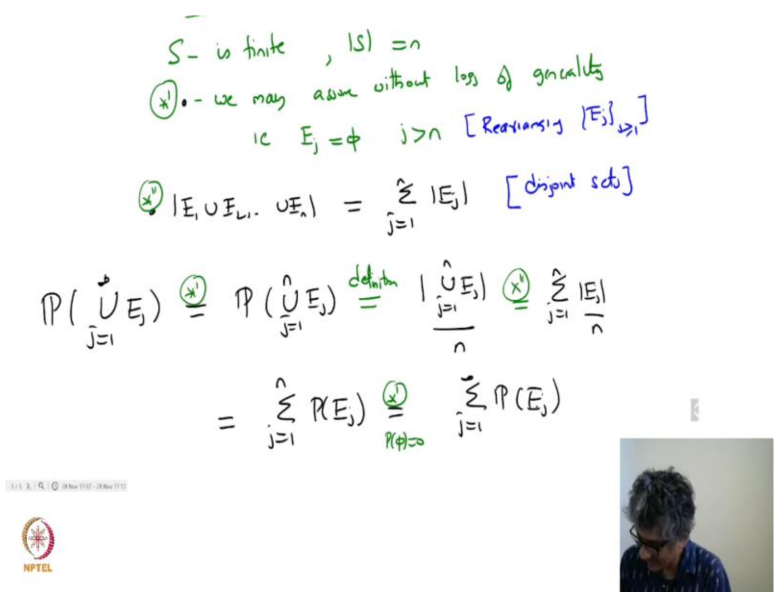
So, for every subset of S, I assign the probability of E to be the number of elements of E divided by n; where n is the total elements in S, number of elements in S. Now, the claim is that this definition automatically ensures two things that P is indeed a probability of (())(10:12). And moreover, P assigns equal probability to every individual outcome; so, let us try and see how to prove this ideally. So, to to show P is a probability on S what I will have to verify, how to verify first. P is indeed a function from 0, 1, 2 P is indeed a function from F to 0, 1; I will verify this right, so is this true? Let us verify this.

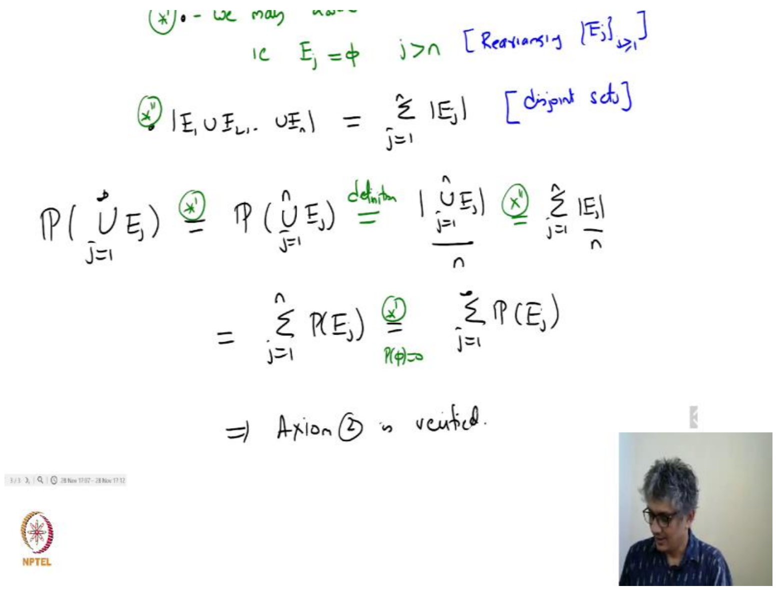
This is true this is indeed true as E is subset of S, will imply to number of elements of E is always non-negative. You see it could be empty or to be 0, is less than or equal to the number of elements of S. So, which implies that by definition P of E, which is equal to mod E by mod S which is equal to n; is in fact an element of 0,1. This is a number between 0 and 1; because if I divide E by S, what mod S will equal to n. So, let me write the mod S will be the same as n. So, if I divide if I divide E by mod S; I get a number between 0 and 1.

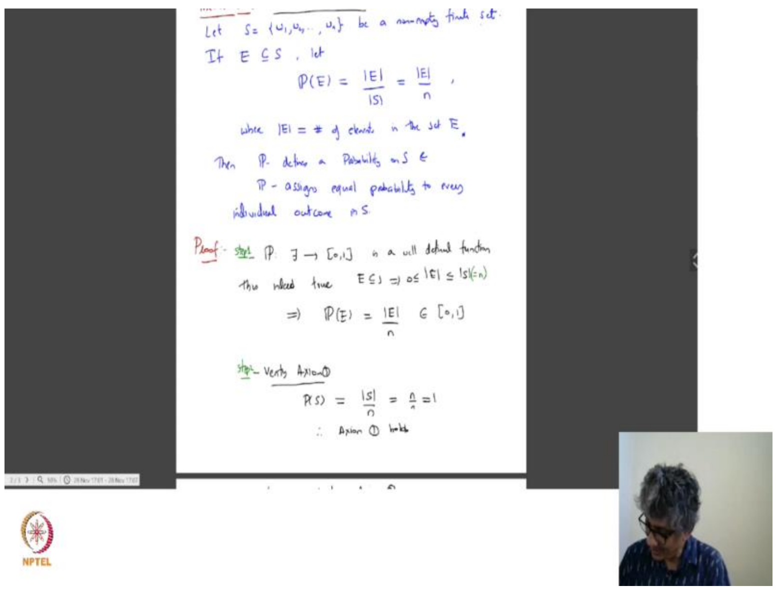
So, P is indeed a good well-defined function; so, P is a well defined. So, the first claim is that P is a well-defined function; the second thing is I have to ensure the two Axioms of the probability. So, let us verify Axiom 1; so, let us call this as step one. Step one is to verify, it is a function; step two is to verify Axiom 1. What is Axiom 1 say? Axiom 1 says the probability of S should be 1; but what is probability of S? Probability of S by definition is mod S by n. Because that is what the definition of any subset S is size of E by size of S; so mod S by n, which is same as n by n which is equal to very good. So, indeed it is very fine, so therefore Axiom 1 holds for this function.

(Refer Slide Time: 12:39)









Step 3, what is step 3? Step 3 is I have to verify Axiom 2. How do I verify Axiom 2? Axiom 2 you can verify the following way. You have to take a countable sequence of events, so let us say E1, E2 and so on and so forth and En are a countable sequence of disjoint events. So, perhaps I should say let let E1, E2, En be a be a countable sequence; so this is given to us. So, now you observe little S is finite, mod s s is finite, s is finite; s is finite and mod s is equal to n, so capital mod s is only n elements. So, in some sense if I have a countable sequence of disjoint events, if I have disjoint events; they have no intersection.

Then after a while they all have to be empty because I would have exhausted all the elements in S. So, this is a small exercise, think about; we may assume without loss of generality that is we can rearrange the this the Ei’s in suitable form generality that. That only the first n elements are non-empty; that is E’s of j is empty for j bigger than n. So, this is the you have to you have to just do it by just rearranging Ej’s; because you can only have at most n disjoint events in S non-empty events. Otherwise, you will have empty ones, and you also observe this following fact.

So, now once you have this, you know that size of E1, E2 union En if they are disjoint, is the same as the sum from j equal to 1 to n modulus of Ej; that is just a fact if the sets are disjoint. If your disjoint events, disjoint sets if you add them up the size adds up; so, if you take the union and take the size the same as adding the size of each. So, let us go to our proof now, what you have to do now? You have to these are the remarks you have. This is one important remark and this is one important. So, now you whatever you do, you take the probability of the union j equal to 1 to infinity E’s of j.

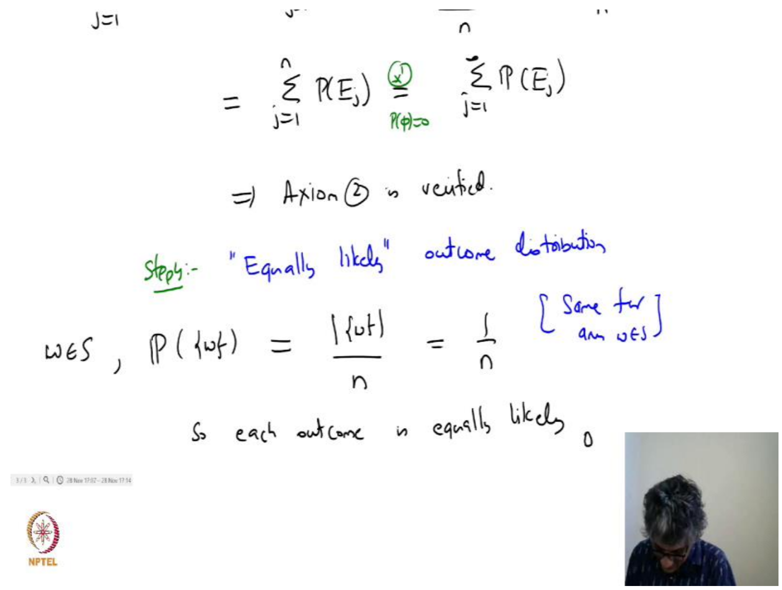
So, what is that going to be? That is going to be again; so E’s of j is empty after n. So, that means this is the same as probability union j equal to 1 to n E’s of j. Let me call this as maybe some some number let say let us call this star; star prime, this is also star prime because after some time everybody here. Now, we know this is by definition by definition, what is it? It is the size of the union j equal to 1 to n E’s of j divided by n. But, by this other fact let us call this star double prime now; so what do I get? Star double prime will imply that I get the sum of the Ej’s. Sum of the size of the Ej’s, modulus of Ej’s divide by n.

But, lo and behold this guy essentially is equal to the sum j equal to 1 to n probability of Ej; but now because probability Ej is just size Ej by n. But now, use the fact that again you start prime again you star prime, and the fact that probability of an empty set is 0. Use both these facts to get this is the same as summation j equal to 1 to infinity probability of Ej.

So, the after n everybody is empty their probability is 0; I can (())(17:27); so, essentially I have verified Axiom 2. This implies Axiom 2 is verified; 2 is verified, that is very good. So, so that means what have I done I verified that this function defined in this manner, is probability of P equal to size of E by size of S is in fact a probability.

That it is well defined, it verifies Axiom 1 and verify the Axiom 2; let me comeback and review the proof in a second. But, what is left? I have to show it assigns equal probability to every outcome; that is quite easy, let us see how this why is that easy. Let me just go back down, so last step is to verify that each let us do step 4.

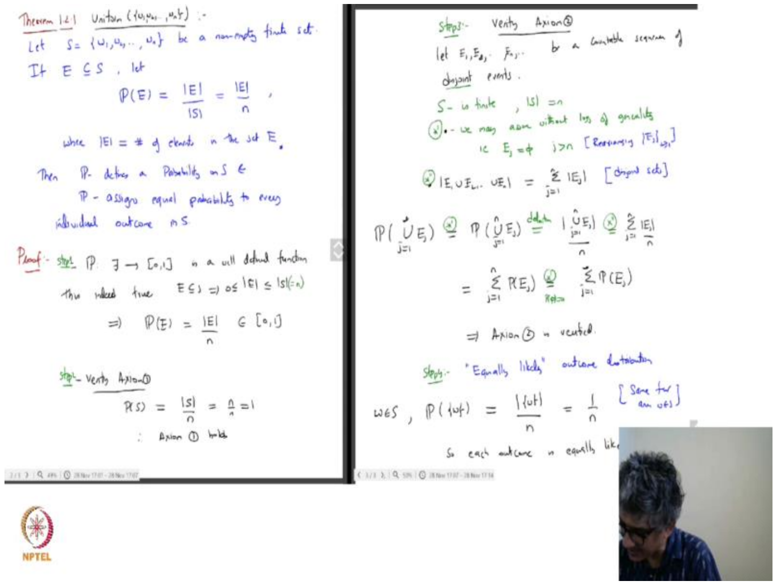
(Refer Slide Time: 18:17)

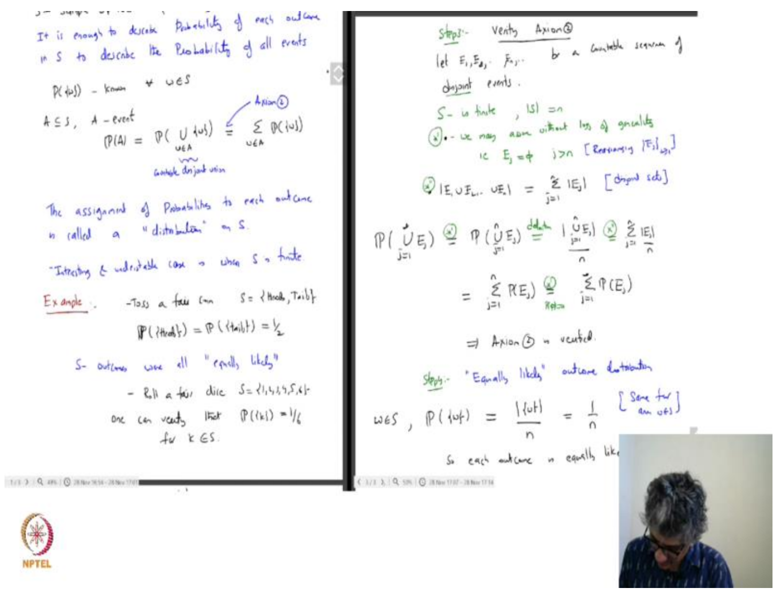


So, I will just call it the equally likely outcome distribution. So, how do I verify this? I have to take the probability of let say omega is in S, some omega is in S; some element in S. Probability of omega by definition is what the size of omega divided by n; but size of omega is just one, so just 1 by n.

So, that means this is the same for every omega, so same for any omega; that means we verify the fact that every outcome has the same probability. So, so each outcome in S has the same probability; but I would write it as each outcome is equally likely; because they all have the probability one over.

(Refer Slide Time: 19:35)





So, it is a very important nice so let us go back, and since I want to stress this on this a little bit; let me go back and redo this proof given in one summary shot. So, the idea was that I wanted to I want to come here whereas if you can step 1, step 2, step 3 very good. So, the idea was that I wanted to imitate this idea that that I have a toss of a coin fair coin; probability that any outcome is equally likely.

Because probably head is half and tail is also half; similarly the role of a fair dice, the chance that each outcome will come is 1 over 6 you can verify that. So, then that motivated us to define a little more general manner; that is called the uniform distribution of n elements, so omega 1 through omega n. So, what you do is you define the size of any element E has a number of elements E in that set. And then you define the probability of an event E has the size E divided by n; and once you done this look at this. I can claim that P defines the probability of S and P assigns equally equal probability to every individual outcome.

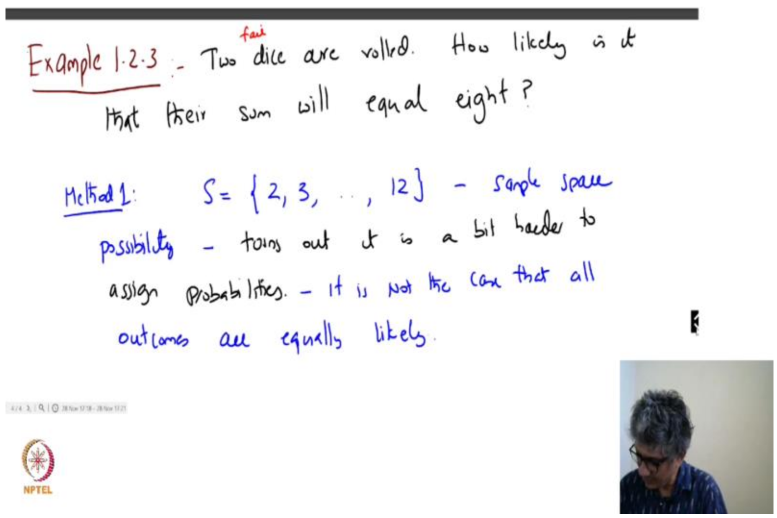
So, further what did you have verify? You have verified that this definition indeed defines the probability. So, what I did was I took the map P from the event space to 0, 1; I had to show it is a well defined function. But this indeed true because you take a subset of E; the number of elements in E is going to be less than number of elements in S. Because the subset of S that means probability of E is probability of size of E by N; which is the number being 0 and 1, so 1 is verified. To verify Axiom Axiom 1 of probability; I have to verify probability of S is equal to 1. But that is not so easy because S probability of S is size of S by n, which is n by n; so, Axiom 1 holds.

Now, verify Axiom 2 the same way that E1 through En be a countable sequence of disjoint events; S is finite, size of S is n. Then the crucial thing is that if you have n elements in S and you have collection of disjoint events; after sometime they have to be empty, so you can rearrange them. So, the first n elements are at most non-empty, but everybody about that is empty; that is what star prime does. Then star double prime just does as a simple fact, if you have n disjoint events; the union size the union is the size of the sum of each each size. That is what second one is.

Once you have this you just have to observe that the countable union first reduces to the finite union; because everybody is empty after some time. The probability of finite union is this is defined as the size of the finite union divided by n. But, that by star double prime is same as the sum of the the size of this prime. And the low and behold size of Ej by n is same as probability of Ej; and that implies that the sum of j equal to one to probability of Ej is the same as the countable set, so Axiom 2 is verified. Once Axiom 2 is verified equally likely outcome distribution is also automatic; because probability of omega is size of omega by n, which is 1 over n.

So, it means I verify both these tables; the P is the probability by the first three steps. And the fourth step tells me that probability of every outcome is 1 over n is the same outcome. It is a very important idea; this is called the uniform distribution on n elements; and the way you define it is that you define it by the size of the set divide by n, that is total number of n elements the sample space. Now, let me do couple of examples to sort of set this notion correctly and I will close this discussion.

(Refer Slide Time: 23:54)

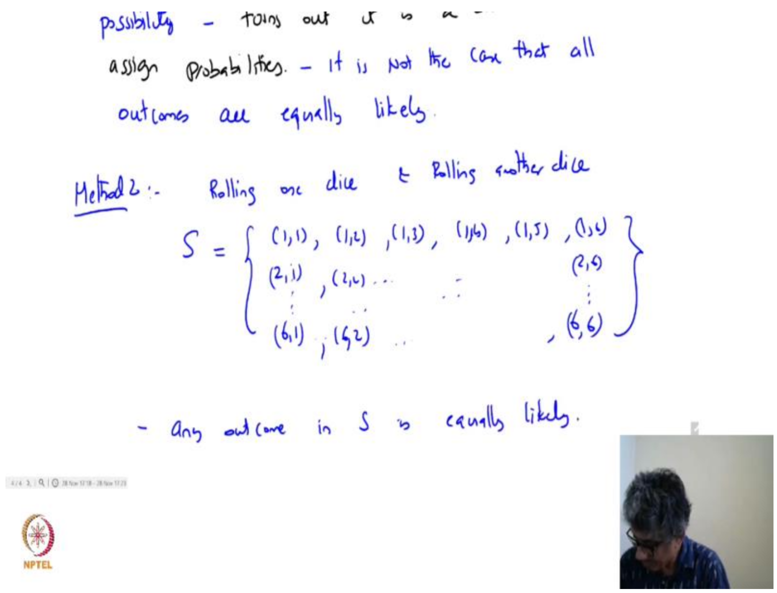


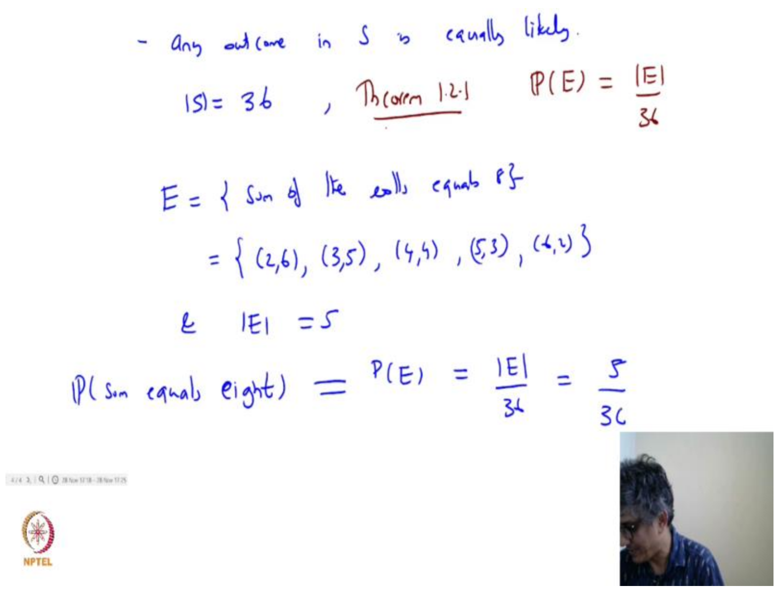
So, here is an example, example 1.2.3; so let say I have two dice, two dice are rolled. And I want to know how likely is their sum sum will equally likely; two dice are rolled. I want to know how likely is it that their sum, let me sorry how likely is it that their sum is equal will equal eight? That way as I used to I rolled two dice and let us assume they are fair die two fair dice are rolled. I assume I want to know so there are various they are setting this up;

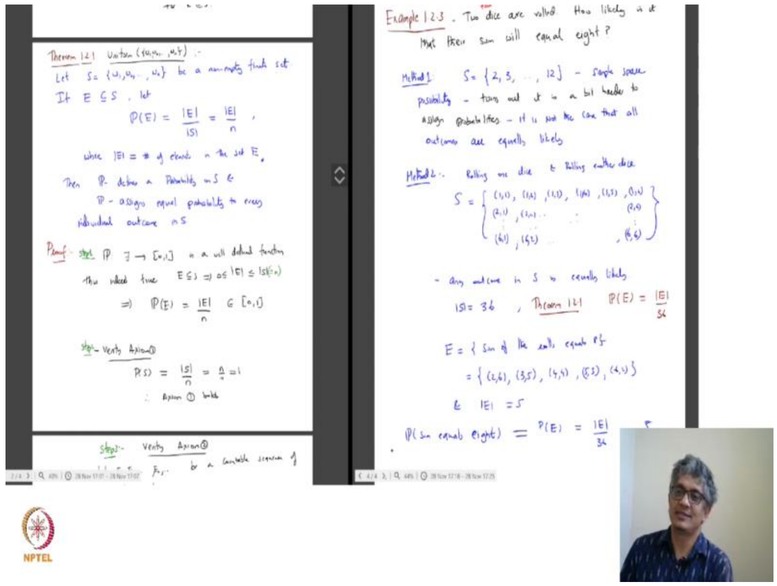
So, let us do let us do method 1. So, method 1 you could you could say fine, I want to I have two roll; I am rolling 2 dice, I am interested in the sum. So, the sample space S will be the sum of the two dice; what are the possibilities of sum? Both are 1, 1 it is 2; 1, 2 is 3 and so on so forth up to both are 6, 6 I will get 12.

So, this is S sample space, this is a sample space possibility; and then you would say it is fine. Now, I would like to compare and find the probability of junk; but this turns out a little hard. It turns out that once you do it like this it turns out; it is a bit harder to assign probabilities. In the sense that you have to think a little bit what are the possibility that will give you three; and how likely each is and so on and so forth. And it is definitely not the case that all outcomes are equally likely; and it is not the case that all outcomes are equally likely. So, instead I I I I do it like this, I think of it I do not do it this method; I abandon this method; I do a next method, method 2.

(Refer Slide Time: 26:53)







So, in this method what I do is I say fine; I do not think of it as observing the sum. I observe it I think of it as rolling one dice rolling other dice; noting the outcome and then I compute the sum that I want. So, what I do is I view the experiment as rolling a dice one dice and rolling another dice. So, then the sample space S becomes, so the first outcome could be 1; the second one could be 1, then 1 2, 1 3, 1 4, 1 5 and then 1 6. Then you go all the way down, you get 6 1 and all the way to down 2 1 and so 6 1 and then 6 1, 6 2 and then 6 6.

And here you have 2 6, 3 6 and so on and so forth; and here you have 2 2 and so on and so forth. So that means you you list out all possible outcomes that can occur in in the two rolls; or one dice and second dice. Now, you know that if you roll a dice each outcome is equally likely; that means any one of these these points; the sample space is equally likely. So, now we are in this so any outcome in S is equally likely; that is something you observe immediately. So, now that means I can use my previous theorem; so that means the size of S is what? Six possibilities in the first dice; six possibilities in the second guy, so size of S is 36.

And equally likely experiment, then we know from this theorem that we did just now this theorem of 1.2.1; that let me write down that in the rules and scales. So, theorem 1.2.1 in the (room) that the size of E is going to be probability is going to be size of E divided by 36; that is what is going to be, very good. So, if you want to define a probability on this S, it has to be the size of E by 36. But now what are we interested in? We are interested in the event that the sum will be equal to 8; let us write the event down. What is the event E? The event E is sum of the rolls equals 8.

But, sum of the rolls equals to 8 means what that means? If the first guy is 1 then I cannot get 8; because the highest is 6 for the second guy 1 plus 6 is 7. So, I will have to start with 2, so first guy is 2; the second guy can be 6 there is no other choice. So, once I get the first guy then S is, the next is the first guy is 3; the second guy has to be 5. The next is 4 4 4 and then I can go bigger, I go 5 and then 3 and then 6 and then 2. So, these are all the outcomes that will give me the sum is 8; let me and also know immediately that the size of E is what is 1 2 3 4 5, so I review as 5.

That means the probability of E is equal to the size of E divided by 36 which is 5. And if I do a little jugglery we move this around so this (())(30:29). I am going to do this at one last time, very good moves across; so, now I will just erase this part right here. So, what I wanted to say was that the probability that the sum equals 8, is the same as the probability of E; the size of E by 36 is 5 by 36. So, you you have so what the point of the exercise is the following that you have to be little bit sort of careful. When you are given a question if you want to use the previous theorem; you have to set up the problem correctly.

If you go method 1, you got a little bit stuck a little bit; I do not know what to do, what is going. I have not assigned probabilities also, but once you set up like an equally likely experiment; then that then the challenge then becomes quite simple. All you have to do is make sure you identify the event E correctly at some of these; so, just quickly recall the whole setup that we did today. One was that the whole point was how to throw a uniform distribution on n elements; and the way one does that is that you assign the probability in the event E, has the size of the event E divided by mod S.

And then that this theorem shows that this defines the probability on E, probability on S and assigns equal probability at the outcome. Then this example tells you that if you want to apply this theorem; you will be a little bit careful. You have to make sure that you set up the experiment correctly. In method 2, I set up so that S has several outcomes, where each outcome came equally likely in my experiment. And that implies that the size of E divided by total number of n elements, determines the probability of E. Once I have that I identified my event E as a sum of rolls of equals 8, which is equal to 6,2,6,3,5,4,4,5, 3 and 6,2. And that give me the (auxiliary). Thanks.