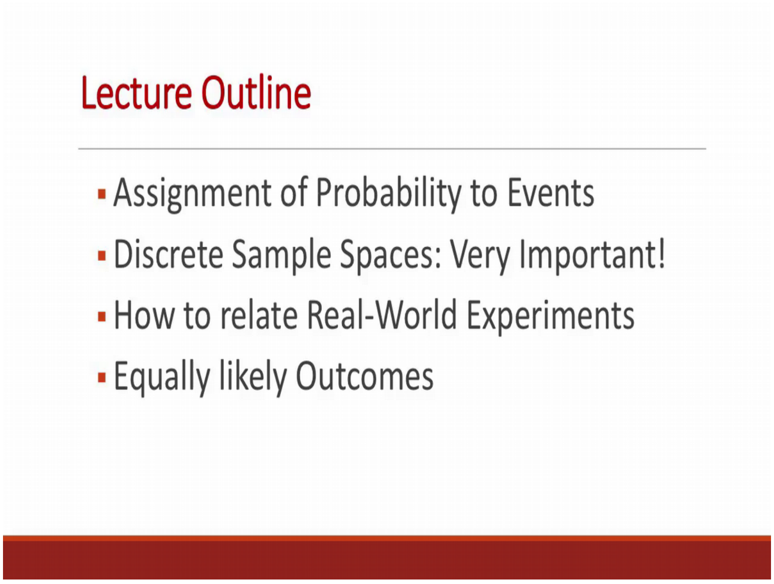
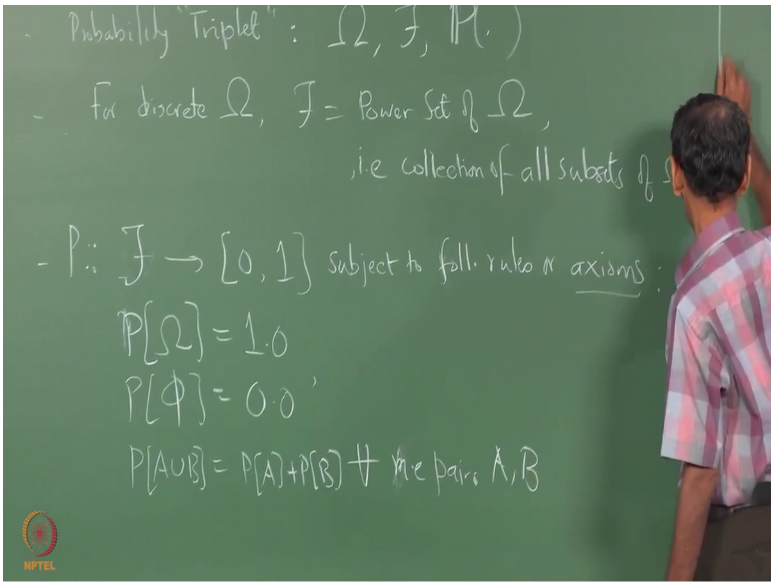
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 04
Axioms of Probability

(Refer Slide Time: 00:15)



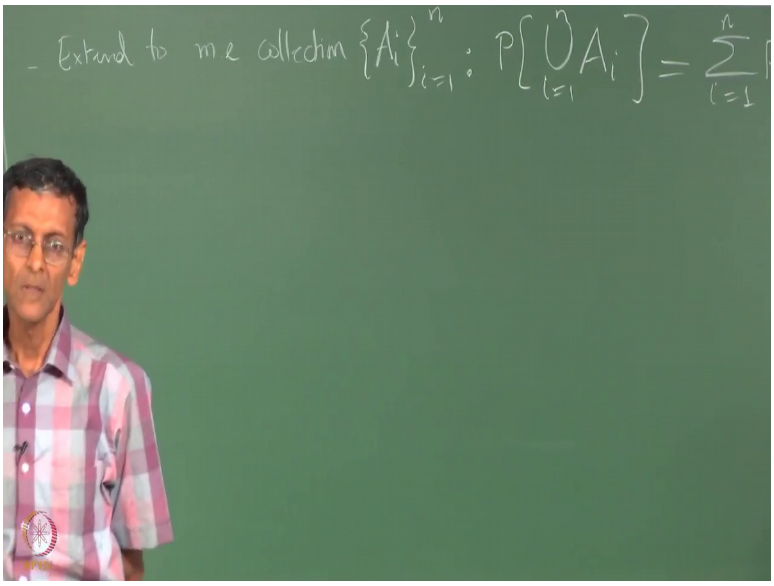
(Refer Slide Time: 00:20)



Essentially the probability function right, takes elements of our members of this collection f, to the real line or the piece of the real line 0 1; that it is a closed interval in the sense that you can get both 0 and 1. In fact, we get both these endpoints right P omega is in fact, exactly equal to 1, and that is a kind of universal convention that we follow right. Everybody uses it worldwide has been using it for the last 15 over ever since. So, this formalism was put in place right. And of course, the null set has zero probability, which is also very intuitive, because a null set corresponds to the impossible event right; you want it to have 0 probability anyhow. Then the most important thing beyond this is this statement here right.

If you take any two subsets A or sets A and B which are mutually exclusive, then they there the probability the union of A and B must be the sum of P A plus P B. So, these are all, the rules or axioms that we use right to construct this probability function right. So, if you have any doubt about what an axiom is, you can think of it as a rule right, you do not question it basically right. So, this pair wise additivity axiom right; obviously, extends to any collection of, a finite collection of mutually exclusive subsets. Note that right.

(Refer Slide Time: 01:51)



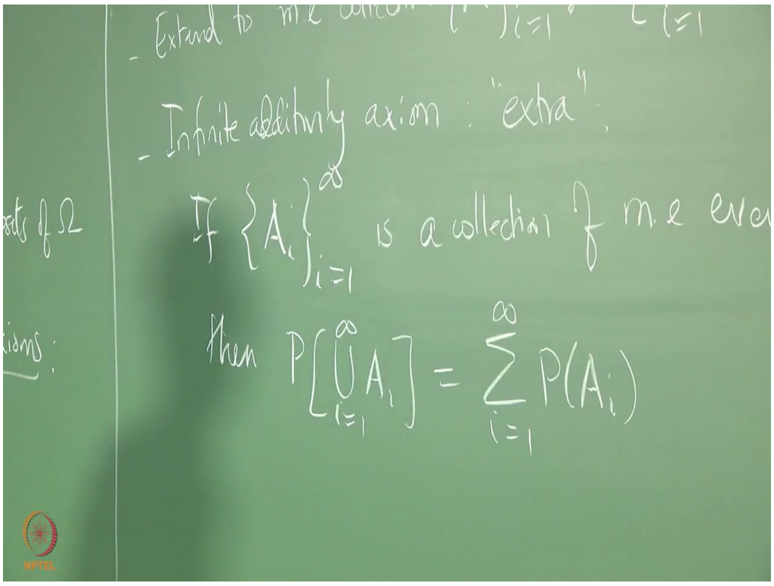
If A B and C are mutually exclusive, any pair wise also they mutually exclusive right. So, you can have more than two mutually exclusive subsets is in a sub collection right.

So, extends to axiom number 3, a mutually exclusive collection A i, i equal to someone to n let us say right. This union right is also; obviously, defined and its set in f, because we have said that, at least in f right unions have to be a members of f right. So, there is. So; obviously, we have to assign a say probability to, the union of any collection of sets which are in f. So, the assumption automatically is that all A i s are in f. Again for the discrete case the all these questions of what is f does not arise, because we are almost. We are always going to take f to be every possible collection of, every possible subset we can form out of omega right; so anyway.

So, all of these things, statements are going to be true for all omega f P right. Regardless of whether f is a power set or not, which is why they are, you know these. This is just an example in a way right. It is a special case, where as all of this applies to everything, or any probabilities are triplet that you define right. So, for any this pair wise at the additivity clearly extends to, any collection of n mutually exclusive objects, because of induction right if A B and C. You start with A B and C, which are all three exclusive of each other in the sense only one can happen at any one time right. Then A union B will be exclusive of C and so on and so forth right.

So; obviously, this extends quite; obviously, what does not follow from this finite additivity axiom, is the infinite additivity, which has to be also added as an extra thing. And that the infinity additive axiom needed, when you go to spaces which have an infinite number of elements right, whether countable or uncountable so right.

(Refer Slide Time: 04:37)

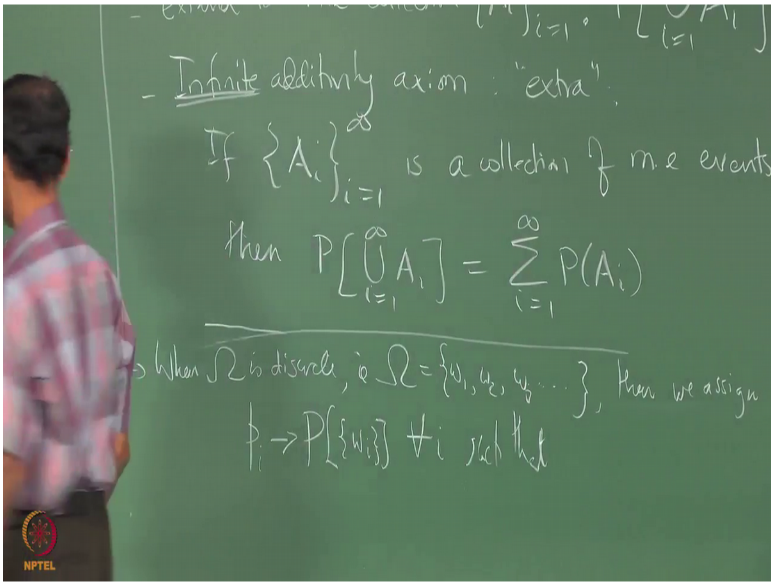


Now so called infinite additivity; this is a extra. This is an extra. So, if this A i want to infinity is right; the collection of mutually exclusive sub events, then so this is stated on top of this right. Note that this again is a number which has to lie between 0 and 1 right. And this has, this union has to be no bigger than omega ok.

So, the infinity really does not, really cause any problems at all right. And; obviously, you will run into this kind of situation when omega is big enough to support right, an infinite collection of subsets right, whether countable or uncountable rising; so anyway. So, with these axioms right, we can now proceed with the development of the theory right. What are the properties, other properties that we can derive from them and so on and so forth right? So, we should take. I mean little bit of a pause to and understand this and appreciate this and let it sing it to our heads right, that these building blocks are in fact, sufficient pretty much for to build up the entire theory right. It is a very profound in a sense, situation right that a very small number of rules and all that we need to keep in mind.

We need to keep in mind how we form omega, and then the f, and then apply this you know the P to all the elements of f, and then follow these axioms to build up the theory right. Now before I move on to the actual, you know the probability is that. Sorry the properties; so of the probability measure P right. Let us take up the example I know the assignment of probabilities to elements in f for the case that f is countable, or sorry when omega is countable right.

(Refer Slide Time: 07:27)

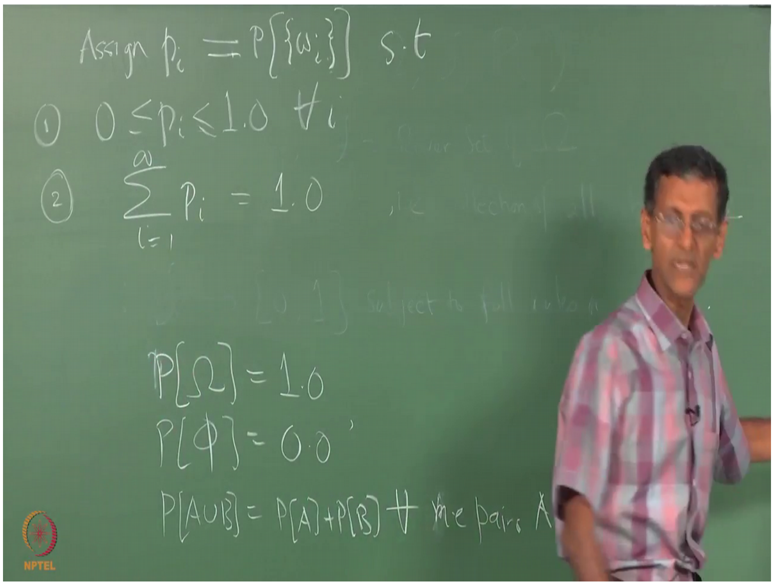


For discrete, when omega is discrete i e omega equal to; capital Omega is Omega 1 omega 2 omega 3 and so on dot dot, without right, can go on forever if necessary, if need be right then right. So, how do you go about assigning piece so that all these axioms are satisfied right?

Basically, you come, you start with a bunch of numbers P i right, what are these numbers P i? Then we assign a number P i to the probability of this singleton set omega i. Some number P i right. What is this notation mean? It is the singleton set right which consists only of the outcome omega i right. So; obviously, a singleton set which consists only of omega i, is exclusive of any other singleton set which consists of different outcome right. So, note there is a certain difference between writing just this outcome omega i, and writing it as a singleton set right. Although in engineering terms there may not be any difference, but in mathematical terms there is a difference right. This is a singleton set, this is just an outcome right anyway.

But you can think of P i is just the probability assigned to omega i right. So, we assigned this P i to this singleton set omega i; such that, what rules are we going to follow.

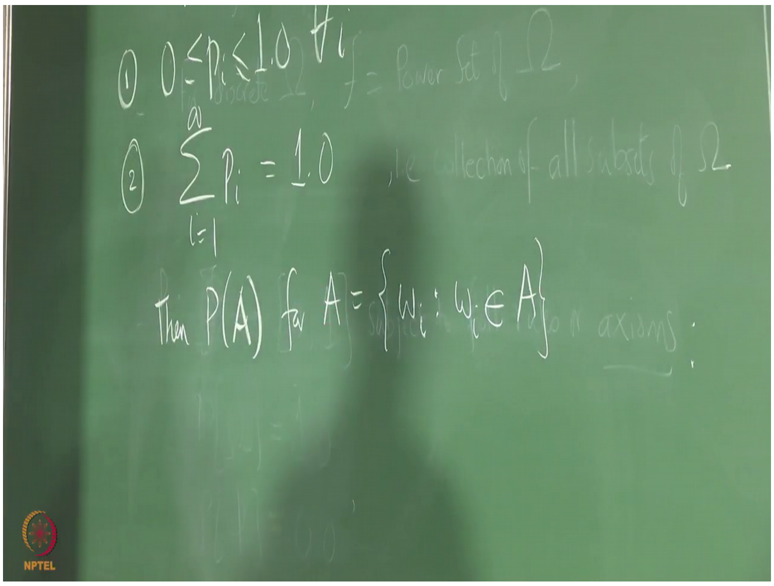
(Refer Slide Time: 09:46)



Let me go back up there and write this right. How are we going to assign these P i s? Let me say P i equal right; such that this P i s is between 0 and 1 for all i; obviously, right this obviously, has to be satisfied. Otherwise P i cannot be a probability right. Then secondly, sum of P i, i equal to 1 to infinity is 1. So, when omega is discrete. All of this applies only to this case, or omega is discrete right. It is very difficult to do probability assignment consistently, if you start with all the subsets to which or all the sets in up to which you want us and probabilities right. It is much easier.

And this is what we have always been doing in all our earlier right encounters with the subject. We have done it implicitly without right putting a formal statement on it; now it is time to look at it a little more formally. All that we have done up to now right when omega is discrete is, assign probabilities to this a singleton subsets; such that the probabilities added to one. The probabilities are between 0 and 1 themselves; so these two things right. The statements are sufficient once again to generate any, the probability of any event which is going to consist of more than one of the elementary outcomes right, of the elements of the outcomes of capital Omega. How, because the simple additivity rule right, then gives you this supposing you want P A, some collection of omega i.

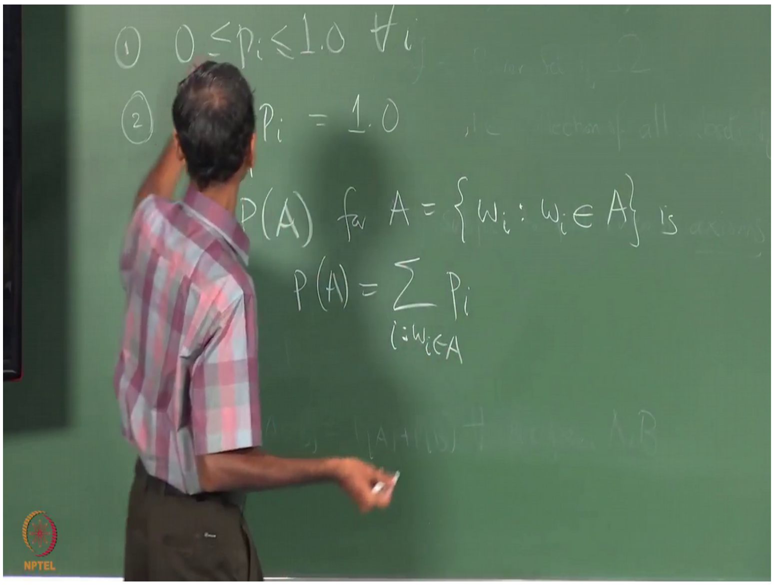
(Refer Slide Time: 11:38)



Such that right. This is a circular definition right, but anyway let me, does not matter I am not being very rigorous in mathematical terms, but. So, let me just write it in some notation like this. This colon stands for such that right. So; obviously, it is you know. It is; obviously, a circular definition, but does not matter let it be right. So, since I do not want to tie down my A to B an example of omega 1 omega 2 or whatever right. Some collection of points omega i is in A.

So, what will be P A, then what is P A in terms of P i, I have put a summation right. i added all the P i s, but; obviously, I do not add all the i s from 1 to infinity, I add only those i s; such that what omega i belongs to A.

(Refer Slide Time: 12:45)



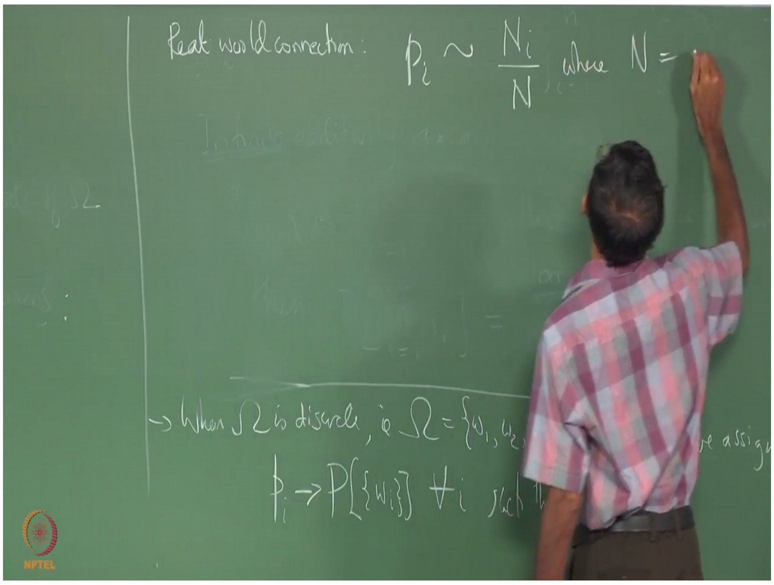
So, this has to be written mathematically like this i; such that omega i belongs to A. So, you apply this simple principle to any set A or any event A right.

So, this itself is consistent with the axioms right. After all what is A. A is the union of all the elementary events right, of corresponding to the omega i s that belong to a right. And that has to be obtained as an addition, according to the actions we have just written down. So, therefore, that addition is here. P A is; obviously, between 0 and 1. Can never be below 0, can be adding non negative quantities. So, the result can never be negative either, result can never exceed 1, because if you put all the P i s together you get unity. So, clearly its between 0 and 1 right. And if you take any two sets A and B; A union B right will consist of a larger collection of points right, larger collection overcomes.

And definitely the probability of A union A union B is going to be the, you know the summation of P A plus P B. I think it is obvious from this construction. I do not think I need to spend more time on that than this right. Just keep adding that is all right, of union corresponds to addition when sets are, or the events are exclusive right. So, this is a very intuitive way of assigning the probabilities, which we start doing right when we start selling the subject right. Note that we are not putting any right. The theory does not say how we should come up with these numbers P i, especially when the number of points is large right, number of outcomes is large or right, or even countable infinite right. It does not say how we should choose a P i s.

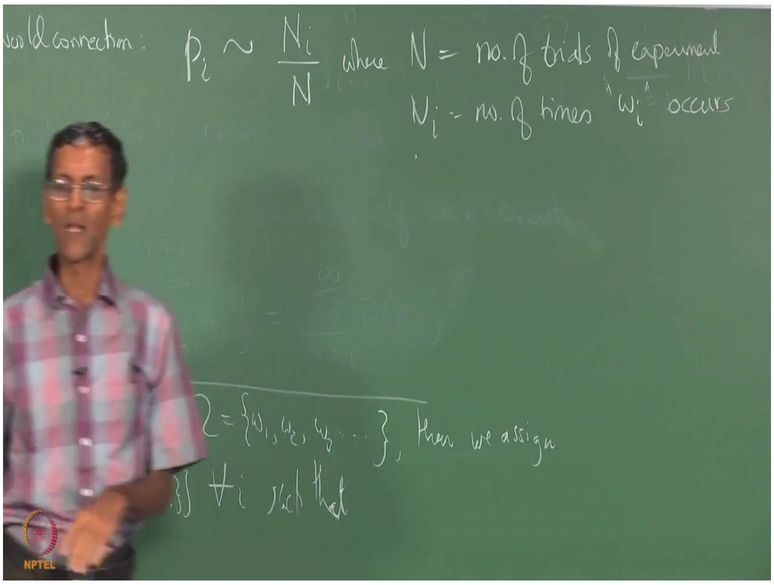
So, then the P i s must reflect reality in some sense right. Otherwise the calculations we do with the P i s, are not going to mean much at all. So, how do we make sure these, how do we get some corresponds to real world, through this concept of relative frequency right. If we run the experiment, a large number of times right, and then count right the number. Let us say we run the experiment some capital N times. I have used small n out here.

(Refer Slide Time: 15:49)



So, I have to use a different N, so some. So, since n is out there, small n. Let me.

(Refer Slide Time: 16:21)



Let me. So, there, what is the real world connection? What is N i? N i omega i; so probabilities are always associated the relative frequencies in application right, but that is the application of the theory right.

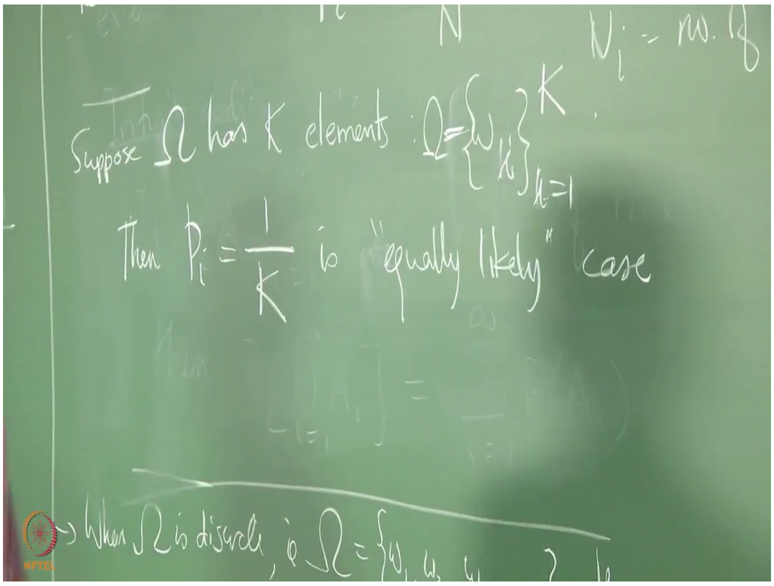
So, the theory comes in between the observation and the application. You observe something, then you build a probability model, you compute using the probability model, and then you have some other numbers that you have to take back to the real world right. So, the theory itself is diverse from both the observation finding and the application. Its completely exist in almost like independent of it the real world right, except that it has to follow the axioms, but for the theory to give you useful numbers say the output right, the input has to be meaningful. The input is meaningful only if, in this discrete case if the P i s right, are roughly right correspond to this N i. I am not putting, note this, I am not putting N equal to, because this N i by N is a very fuzzy quantity right there is no fixed value for different Ns, you make it different N i s, you will get different N i s.

But you may get different N i by N. Is not it. The ratio N i N i by N need not be the same for any collection of N trials right, any set of, any super experiment where you try to do this exponent N times. So, since this can vary, but it is observed right experimentally that right; for many experiments which are repeatable that the N i by N does in fact, right, converge to some number which is reasonably right the same. So, you have to do some juggling of what of the observations, you get to fit into a probability model. So, in this class in this course we are not concerned with that problem at all right. What is a good probability model to use for a particular experiment, we are not going to worry. Its completely beyond the scope of this course right.

That is in fact, a much harder problem, then right, it is an empirical problem. Here we are only building the tools of the theory which are right once. Occasion we use as you know look at applications, but certainly were not going to go out and collect data and try to fit probability models; that is I said it is not an easy thing to do and right, its more in some cases an art, not really a science right. Anyway in some of the cases this P i can be assigned using some symmetry arguments right. If you have a fair coin, and you can only get head or tails out of it and you say the omega is just head tail right. There is no reason to say that probability of head should be much, different from the probability of at A.

So, the standard thing do say yes, they both right have half; half that is very trivial. What about the next most commonly cited example of a die with 6 phase die? Again the 1 by 6 probability associated with each phase of the dice, is just a statement that you do not expect any one phase to occur more than any other phase. All of them are equally likely. So, if unfortunately did, this I did not. Let me use a different notation for.

(Refer Slide Time: 20:10)



Let us say omega; supposing omega has. This is all by the way only for discrete spaces right. We are only looking at discrete spaces now right. So, supposing as they say K elements. This capital Omega is basically Omega K from, or Omega i. I keep this is i, and omega to K right. This assigned, this is just one particular, is what. What do we call this? This is equally likely assignment right.

If you said this P i equal to 1 by K for the right; K as omega, capital Omega is exactly K elements is this so very popularly right thought of as the only. Earlier they used to this is only way that you could do it, but now we know that is; obviously, not necessary right. 200 years back when you know gambling houses, when they started to apply rudimentary probability, they had this right idea, that all basic. The outcome the elementary outcomes from all be equally likely, and then of course, you could get different P As by adding different numbers of outcomes right, but essentially they started off with this right. This is a so called classical theory which was prevalent long back, before axiomatic treatment took over right.

So, this is just one special case right, but it need not be always like this. You could right have even in the case of equally, I mean K, this finite number of outcomes you can have unequal. You start off with an unequal probability distribution. There is no, not the theory, does not preclude that at all right. So, this anyway so having said this right now I think we can move on to the general situation right; omega f. I go back to the omega P, the general case right, which applies in all probability problems or, and look at the probabilities and the properties of probability measure.