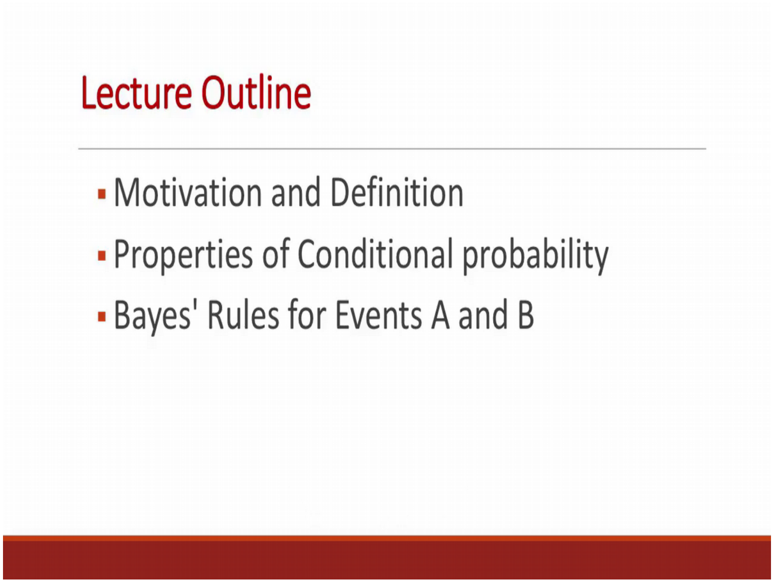
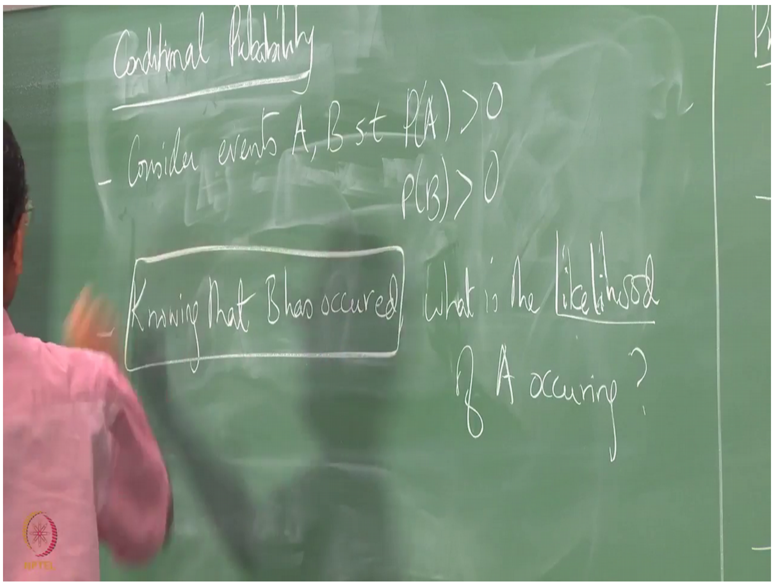
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 09
Definition and Basic Properties

(Refer Slide Time: 00:14)



(Refer Slide Time: 00:18)



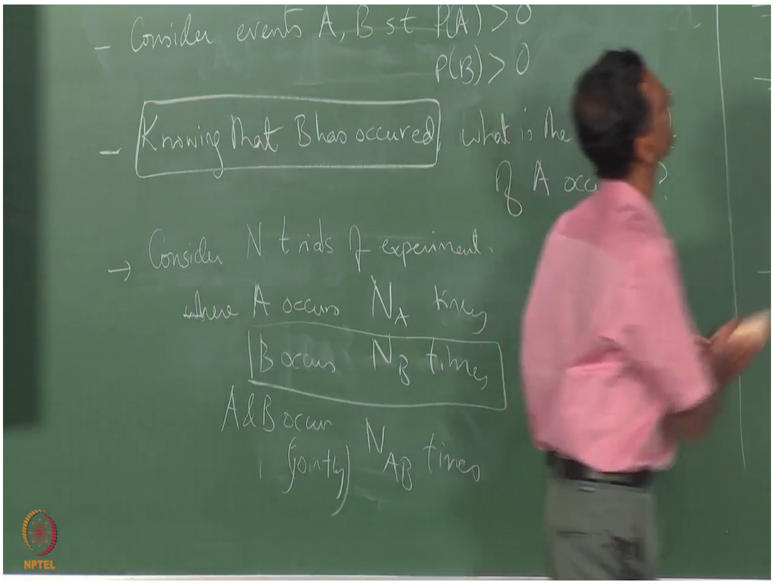
The next important topic that we will look at is conditional probability. Once again examples we will consider collectively right. So, we are just going through the theory.

So, here let us say consider some events A and B both P and, are bigger than 0. This is a very important requirement here because you know you can have events. Not they can be in some case some cases they arise which have essentially end up having 0 probability even though they are not really the null not event. So, we are going to put the condition away upfront that P A and P B are obviously, they are not the null event or the impossible event. So, their probably and therefore, the probably can and it is greater than 0 that probabilities are greater than 0 right. So, that if you do an experiment, repeat an experiment a large number of times you do observe both A and B right.

So, the question is suppose if somebody told you the event B has occurred how does that change your perception about A, you know does it increase the likelihood of A, does it decrease you can it and turns out that an the answer can be both ways. If you know that B has occurred then the likelihood of A occurring or not occurring is can be very significantly different right. So, knowing what is the likelihood? Now this likelihood is another term for probability we are just using different words to in English that are that you will see a l i k e l i, likelihood a occurring or having occurred right.

So, this again is a very important condition that arises often in practice. When you know that something has happened some there is less uncertainty in some sense that you know that B has in fact, happened. So, the third extent there is less uncertainty about the whole situation and it obviously changes or it has a potential to change your view of A.

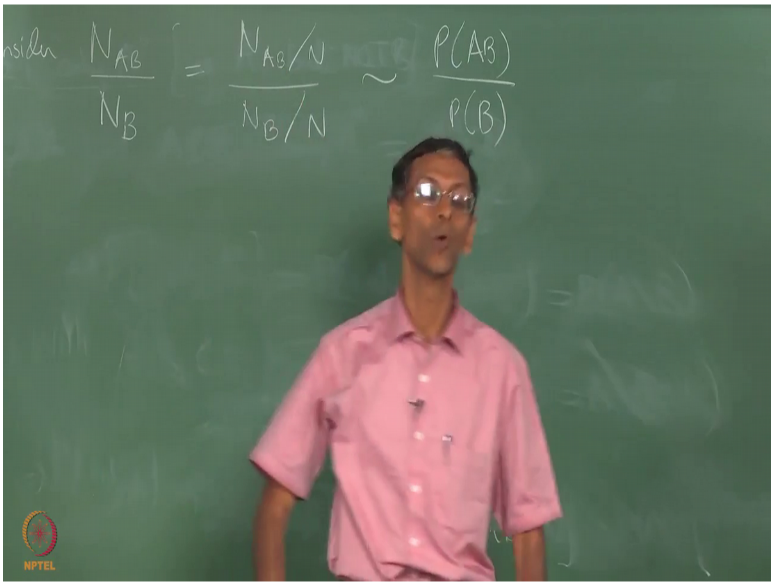
(Refer Slide Time: 03:30)



For example, now if you look at N trials. Now again I am going to use I believe yes in some previous class we use capital N for number of trials, did we I want to stick to that notation. So, there is no confusion we use capital N. So, can if you say if you can say N trials of some experiment and in this N trials let us say that where A occurs N A times, B occurs N B times and one more important thing the joint event A B occurs N A B times which can be 0, if A and B are exclusive which is ok, there is nothing wrong in also including a pair of exclusive events out here right. So, A and B occur this is joint jointly they occur jointly N A B times. So, let us say you have collected this data right.

So, you are not interested anymore in let us say the absolute probability of A. You want to me look at given B means what you are only going to focus your attention on those N B occurrences of B. So, in this subsequence it turns out that N A B is a number of times A also occurs, so N A B is clearly upper bounded by N B, this is obviously, there is no way that N A B can exceed N B.

(Refer Slide Time: 05:43)

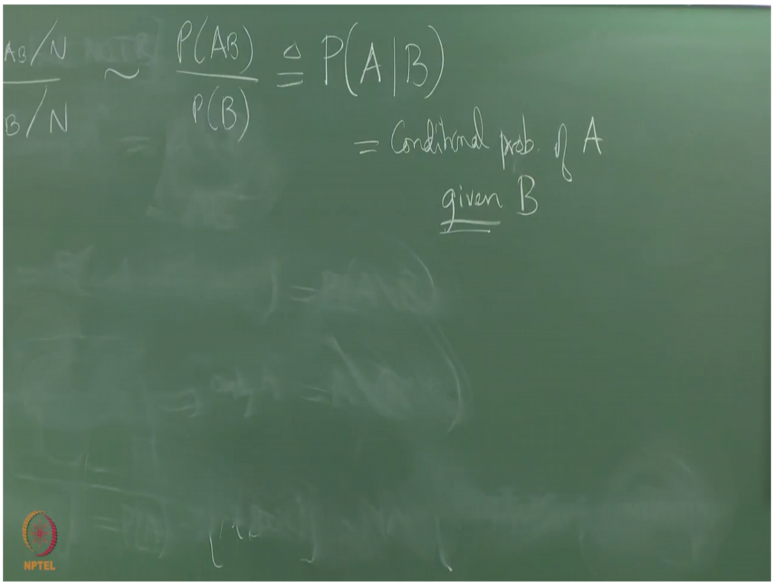


So, now you consider the ratio N A B by N B which is the ratio of the number of times you observe both A and B as well as I mean divided by the number of times you observe only be sorry you observe B. This N B is without regard to A, if B occurs then you say that you count B without any reference to A.

So, this ratio is exactly equal to N A B divided by N and N B divide, the whole thing divided by N B divide N B by N. I am dividing by N in both. Now, if you say if N is large right, this ratio should be reflected in the probabilistic model by the ratio of 2 probabilities which what 2 probabilities are there, is the joint probability of A B divided by the joint, divided by the individual probability of B correct. And we know that we are dividing by a nonzero quantity. So, this here, this ratio from the probability model what does it aims to capture this ratio which is exactly equal to this? So, this is this basis for defining this ratio as what as a conditional probability of A given that B has occurred. So, this is how our likelihood of A changes knowing that B has occurred right.

So, earlier you may have just seen the definition, but why are we using that definition. I think you should ponder about it little bit and realized that this is nothing, but this which is very trivial to actually see, but anyhow. So, but definitely there is some intuition in writing all this. Note that I am we cannot put equals here this is the theory model for this ratio and; obviously, the two are never going to be exactly equal, anyhow.

(Refer Slide Time: 08:01)



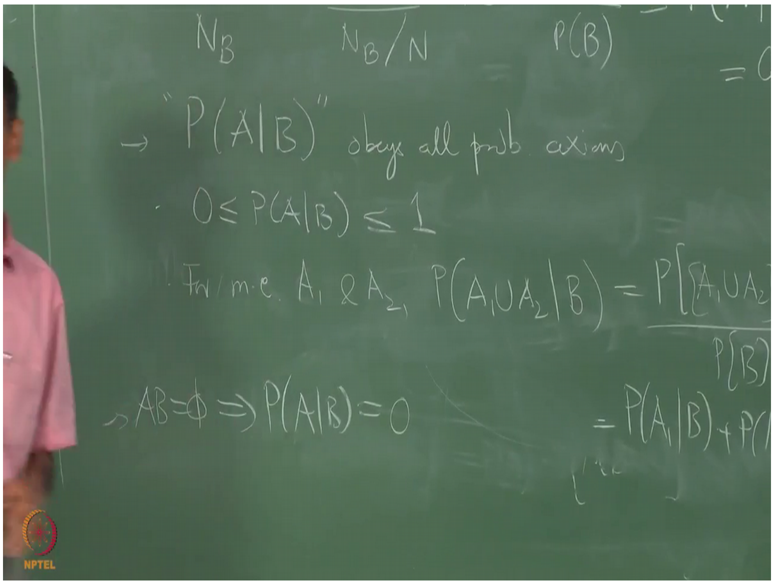
This is written by definition sometimes, this we are not so strict about putting the quantity being defined on the side left hand side and then it is definition on the right. So, basically I am saying that this is being defined by this we defined as this ratio. What is this? This is the conditional probability of A given B. Please read this is one, right; one sentence.

Now, some word on the way in which I have written. This conditional probabilities are always be written with a vertical bar, please do not use write any slanted bars in doing in writing the conditional probability or every conditioning that we encounter in this course and in practice later on should always be written using only a vertical bar. And A and B should come on the same line when you types set, yes if you ever have to I am sure you will have if you write reports you have to do some type setting right. So, you do not have to worry about putting A above and B below or any such thing. Just put it all on the same straight line, horizontal line. And there is no problem, I hope with that getting a vertical bar it is available in you know, it is a basic keystroke right. So, please remember use that vertical bar.

Do not again, when last time when we say right this is a mistake or not, I will not say it is a mistake it is I find this formatting. So, prevalent in usage that I want to say something now, you do not write this out in especially if you in abroad if you go and you do not write it as a and then put over some slanting bar and put B below or something or even to even make it into a horizontal bar A divided by B kind of thing which is meaningless. Some people go as far as that. So, it should never be written like that right. So, I just wanted to do say that very clearly. The first occurrence of this has to be stressed that it is very easily written as just with the vertical bar.

So, this ratio is defined this and then now you find that this P A B by P B the ratio of obeys all the rules that we expect of other probability measures right. So, P A B by P B are; P or let me call it just P of A given B it or I should say P of dot given B, it obeys all axioms when I say P of A given B I am talking of any, P of any event given B you have; obviously.

(Refer Slide Time: 10:55)



Not just one conditional probability you have many condition probability depending on how many events that you are looking at, I know and as we said we have you have a large omega and you take the f to be the power set of omega then this quite a lot of you know (Refer Time: 11:35), you can do a conditioning on pretty much any event here with we are putting B as a conditioning event right. So, you can find define a whole collection of conditional probabilities.

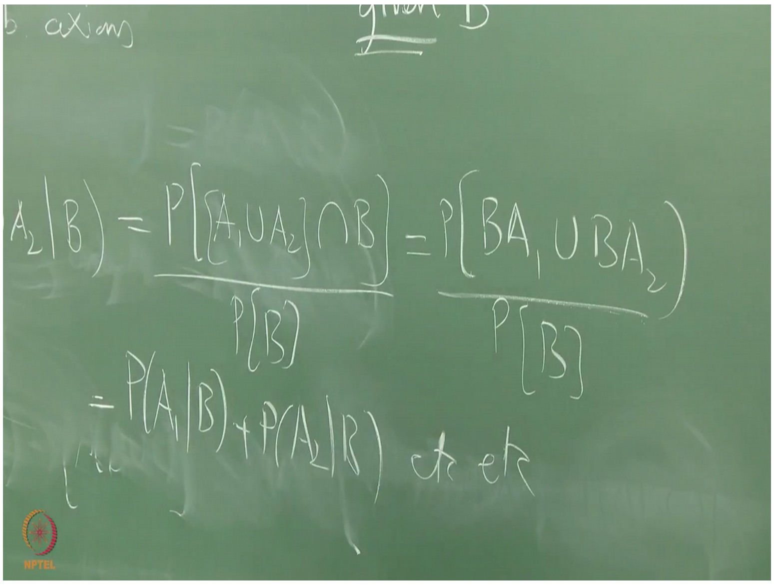
So, what do I mean by this? So, clearly this ratio is between 0 and 1, can never be less than 0 can it be and never be more than 1. When is it equal to 1? When A is in fact, equal to B, when A is equal to B it becomes equal to 1. When A is bigger than B what happens does it exceed 1? No, it will never exceed 1 because A B can never exceed B even if A is bigger than B. For example, if we take A to be omega itself, omega may be bigger than B, but omega B is omega intersection B is B only right. So, this ratio never exceeds B sorry never exceeds 1. So, it is between 0 and 1 and then the additivity axiom is also automatically obeyed because if A 1 and B 2 are mutually exclusive A 1 B and A 2 B are also mutually exclusive.

Let me use A 1 and A 2 rather than A and C right. So, A 1 union A 2 given B if somebody wants to look at this, this is basically P of, I mean what I am going to write here the first step is for any A 1 A 2 not necessarily just mutually exclusive A 1 and A 2, you have P of

A 1 union A 2 intersection B divided by P B by definition right. So, this is true this part is true for any A 1 and A 2 not just mutually exclusive.

So, this of course, can be now this is simplified this is just again this, now we take the B inside so this B A 1 union B A 2 divided by P B, up to here it is there is no change.

(Refer Slide Time: 13:51)



Now, you bring in the fact that let us say if A 1 and A 2 are exclusive, it is joint and exclusive then what happens B A 1, B A 2 are exclusive. So, what happens here? This becomes the sum of B right. So, finally, what will we get in the end you will get P of A 1 given B plus do you want you or do you not get this, you got it right. So, you can extend this very easily to any countable collection of events which are pairwise mutually exclusive. So, therefore, P of A given B obeys all the probability axioms so in fact, you get a new collection of probabilities now all conditioned on B, is not it.

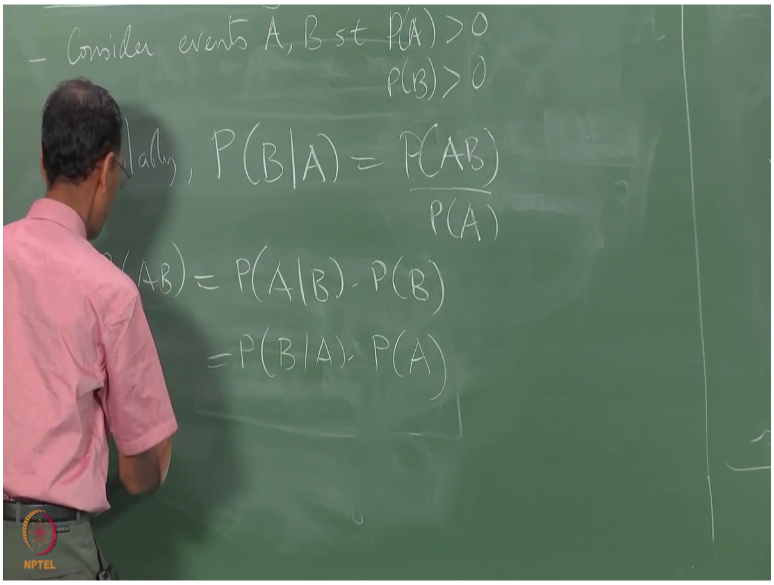
So, I will just write here etcetera, etcetera and you can fill in the blanks if you choose to that is and write it out for arbitrary collection of mutually exclusive pairs of events right. So, one other point here is supposing A intersection B or A B is phi, that is A B themselves are mutually exclusive then what happens this conditional probability will be 0. Let me write it this way A B equal to phi implies that P of A given B is 0. This again makes sense because if A and B are exclusive, if B if you know that B is occurred then obviously, A cannot occur right. So, this is one, there are 2 ends of this when is it 0 when is it 1, A given B is 1 means what? A is at least as big as B, I mean in the sense of inclusion as a set.

The other end is 0 means and we are receiving saying that B P A is greater than 0. So, if P A is greater than 0 A cannot be phi itself, but A B can be phi. So, that is when this becomes 0. So, there is no problem in defining P of A given before any A in the set of sets in the collection of sets that we have in the sigma algebra or sigma field right.

So, we really do have a completely consistent set of probabilities now all conditioned on B all right. So, I am sure this is not the first time this is I am just going over this very quickly because I just this I want to lay the groundwork again, without spending too much time on discussing. This is very important as you all know conditional probability is one of the major additions to the basic theory and has tremendous practical significance. So, I cannot underestimate the importance of this but I do not want to spend more time in that is because I am sure you have seen examples of this already in use, in your earlier exposure to the subject. So, we will keep going unless there is a question. Is there a question? Nobody is raising their hand.

So, one more mean, obviously, before I move on let me also say one of the thing there is no reason why I should only condition on B, I can create another universe by conditioning on A for example, right. So, what happens if I condition on A? Let me start that here and leave this. Now, I am going to say well remove the sentence here, actually all this can be removed right.

(Refer Slide Time: 17:57)

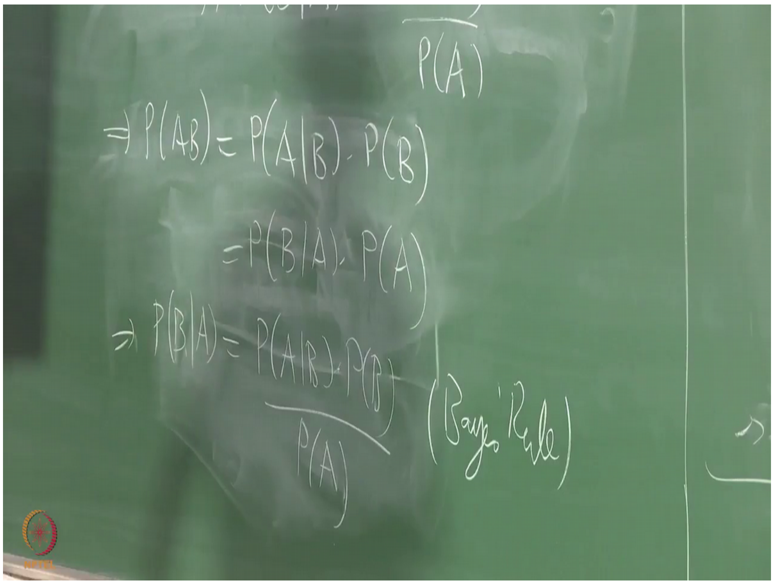


So, similarly P of B given A is clearly P A B divided by P A and here we have no difficulty whatsoever dividing by P A because we have assumed that P A is bigger than 0.

Note that the numerator here is identical, this is the same P A B which occurs in both places right. So, given that numerator P A B whether you get P be B given A or a given B it all depends on what you are dividing by that is all and both cases, both the this as well as this both of them are exactly the same in the mathematical sense they are both conditional probabilities and they have all they have to obey you know follow similar rules right.

What is interesting is the linkage between this and this, that is one form of Baye's theorem that I am sure all of you have seen earlier the fact that they are the same numerator right. So, which implies that this joint event P A B can be written in 2 different ways you can write this as P A B times P B, can also write this as P B given A times P A. Both of these will be important again in our later study. I mean this kind of construction of the joint event. So, this is a nice, is a easy way of thinking about this joint probability, especially in A 2 stage experiment we should be coming to later on where B might be the first stage or an A might be the second stage or vice versa whatever. And it also sets the stage for what we call posterior probability calculation where you know the output of the second stage, but you do not know what has happened in the first stage, the first stage is hidden, but you can observe the second stage right.

(Refer Slide Time: 20:20)



So, for all of these things this is kind of this formulation is going to become very important and clearly we have the Baye's rule which let me just say if you for example, if you want to write A given B or let me put it as B given A here. So, B given A if you want to write this is clearly P of A given B multiplied by P B divided by P A. Now it seem, its.