

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

Lecture - 07
Conditional Probability For Partitions

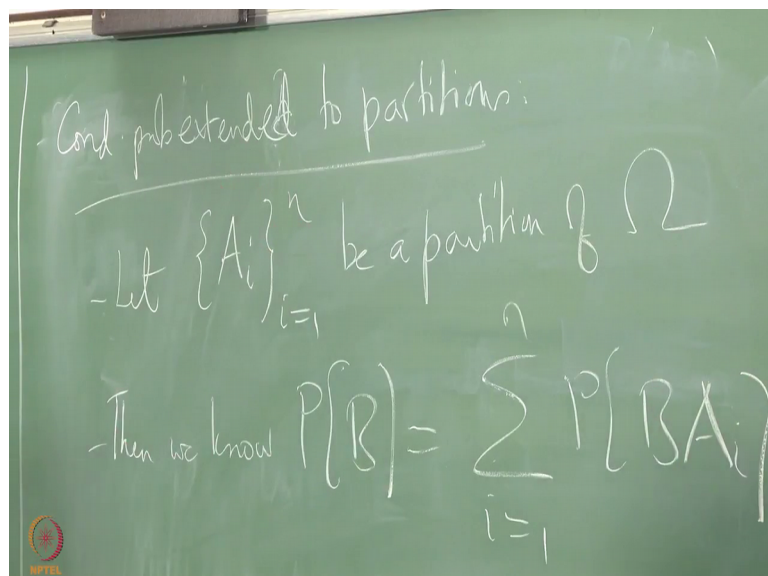
(Refer Slide Time: 00:13)

Lecture Outline

- More general form of Bayes' Rule
- Prior and posterior probabilities of events

So, now can we extend this Bayes' rule to the partition which is again a very important extension right; the whole concept of conditional probability in extent and so on so forth.

(Refer Slide Time: 00:31)



So, the conditional probability extended the partition the const ok. So, here we will say that as we said at the start of this lecture that A_i is a part, right they said A_i collection; A_i is a partition of ω right, then we know $P(B)$ will basically be a sigma of one to $P(B|A_i)$; this is exactly what I wrote earlier. Now, I want to write these guys using conditional probability right, it makes sense to condition on. In this case, let say A_i because they are it turns out in most cases this A_i the partition on ω ; A_i is the first stage and be something you observe in the second stage.

(Refer Slide Time: 01:47)

The image shows a chalkboard with the following handwritten text and equations:

- At the top, it says: $- \text{Let } \{A_i\}_{i=1}^n$
- Below that, it says: $- \text{Then we know } P(B) = \sum_{i=1}^n P(B|A_i)$
- Below that, it says: $= \sum_{i=1}^n P(B|A_i) \cdot P(A_i)$
- At the bottom, it shows the derivation of Bayes' rule: $\Rightarrow P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)}$

So, this becomes using that here this is P of B given A_i into P of A_i $P(B|A_i)$ times P of A_i . So, sometimes this is also refer to as a total probability theorem right where you have the priors, right, the prior probability is in some posterior probabilities and so on. So, we have to be a little clear about the terminology here, right, let me see if I have well I am even in. So, let us just do the match; first right the terminology I will; I have even I am not, let me brush; it up before I come and say here right the what is the mathematics say about this about for example, P of A_i given B that is clearly a posterior probability right that is you know you observe that B has occurred you want to be calculate the P of A_i given B that you can do using Bayes' rule.

So, clearly now the P of A_i is are called the prior probabilities that is your assessment of the situation before making any observation right the P of A_i the probabilities that you have before any right with the old in any additional information is other priors the P of A_i

given B is the all of them they become the posterior probabilities right. So, let me write it out and then. So, what is P of A_i given B this is obviously P of A_i B divided by P B is it not; by definition right and of course, you assume that there is no problem dividing by P B; right then going back here.

(Refer Slide Time: 03:56)

$$P(A_i|B) = \frac{P(A_i|B)P(B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{P(B)}$$

$$= \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^n P(B|A_k)P(A_k)} \quad \forall i \in \{1, \dots, n\}$$

Cond

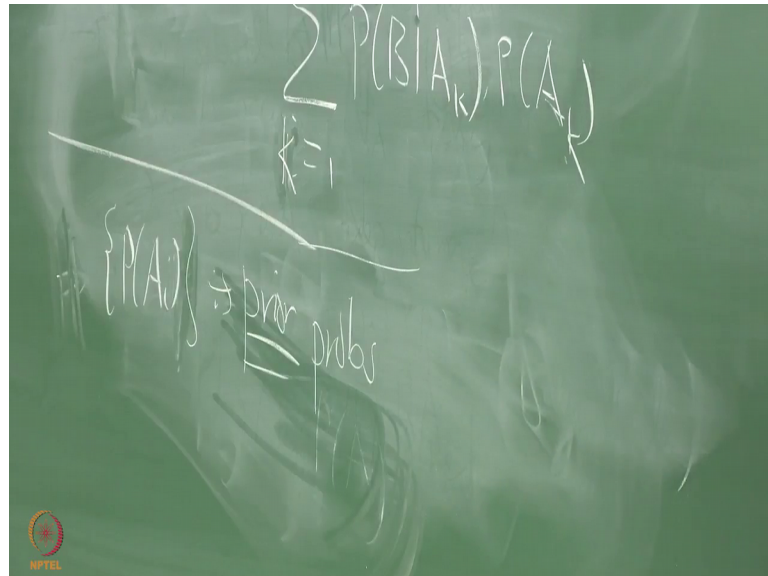
This P of A_i given B let me write this thing here again. So, what will Bayes' rule tell me; it is going to be; clearly if I want to express everything in terms of the prior probabilities in the conditionals P of B given A_i or should I do; I have to write this as P of B given A_i into P of A_i divided by P of B this is one step, then of course, you can go to another step also this will be P of B given A_i P of A_i divided by write this expression for P B is sometimes it is conventional to write down; this all this now you have to be careful in notation right we have already used i here.

So, we should not reuse i; this i we cannot change because that is what right is running through the whole you know left LHS and RHS. So, we have to change this i. So, we will change it to let us say K; K equal to 1 to n P of B given a K and applied by P of a k. So, this scope of this integer K is only within this summation right d the denominator summation. So, there is no K coming outside anyway. So, this is true for all i in 1 to n.

So, for any i; you pick 1, 2, 3 whatever up to n you can express this conditional probability in like this. So, this is saying how the posterior probability. So, this is a

posterior probability right what is right they changed likelihoods of all these events A_i right knowing that some B s occurred again we will come back to examples later on.

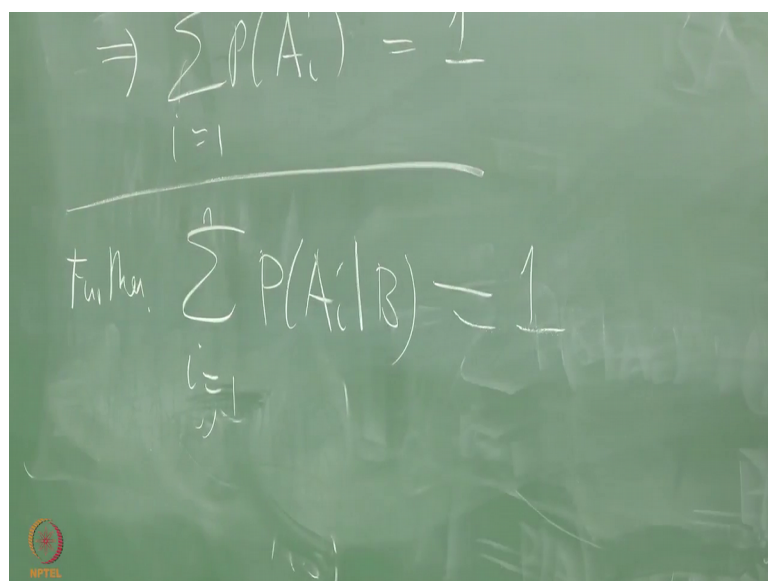
(Refer Slide Time: 06:14)



A chalkboard with a green background. At the top, the formula $\sum_{k=1}^n P(B|A_k)P(A_k)$ is written in white chalk. Below it, a horizontal line is drawn. Under the line, the text $\{P(A_i)\} \Rightarrow$ prior probs is written in white chalk. In the bottom left corner, there is a small red circular logo with the text "NPTEL" below it.

I want to play put do all this first and then examples will take up next week right. So, these this is the posterior probability and I said let me just add the point here this P of A_i is all of them are they the prior probabilities just to just for terminology; I am clearly right because it is a partition of omega this summation of P of A_i has to be unity.

(Refer Slide Time: 06:59)



A chalkboard with a green background. At the top, the formula $\Rightarrow \sum_{i=1}^n P(A_i) = 1$ is written in white chalk. Below it, a horizontal line is drawn. Under the line, the text "Further" is written, followed by the formula $\sum_{i=1}^n P(A_i|B) = 1$. In the bottom left corner, there is a small red circular logo with the text "NPTEL" below it.

So, if A_i is a partition of ω we continue with this point right this implies of course, that $\sum P(A_i)$ the prior probabilities all add up to 1; there is no question about it what about the posterior probabilities $P(A_i | B)$ they also have to add up to one is that obvious or not that is obvious from here right supposing I add up all these guys from one to n it will it not equal this has to equal this right. So, this is this follows from the original.

You know set of axioms we have further here I will use i , right because there is no problem this also this is also equal to 1, it is just what you expect anyway right given that B has occurred 1 of the A_i 's must have occurred because the A_i is a partition of ω always in any run of the experiment one of the A_i 's occurs it has to occur right. So, that does not change whether you observe B or not. So, this is like a sanity check on this whole enterprise right. So, you have prior probabilities right which are independent of any observation right again the wording is very important right you can write or I know pages and pages of match, but ultimately it is your understanding or all this maths is what is bound by the English right. So, do not underestimate the importance of the wording to explain the maths.

So, what are we saying here? We are saying that we have a set of probabilities A_i which are our first guess in a in some sense or first estimate of likelihood of those A_i 's right before any observation is made. Then you make the observation B that can significantly change or likely I mean the way we think of the A_i 's. In particular right if some of these A_i 's do not intersect B at all then what happens; the posterior prior probability can be some positive quantity of the posterior probability will be 0.

So, it is important to always be able to look at the significance of verbal significance of what we have written out using mathematical notation. And that is especially true in this the subject of probability, because these quantities probabilities are very you know sometimes a fuzzy to understand; they are sort of slippery things.

So, at each point I will try my best to make sure I bring out the English also.