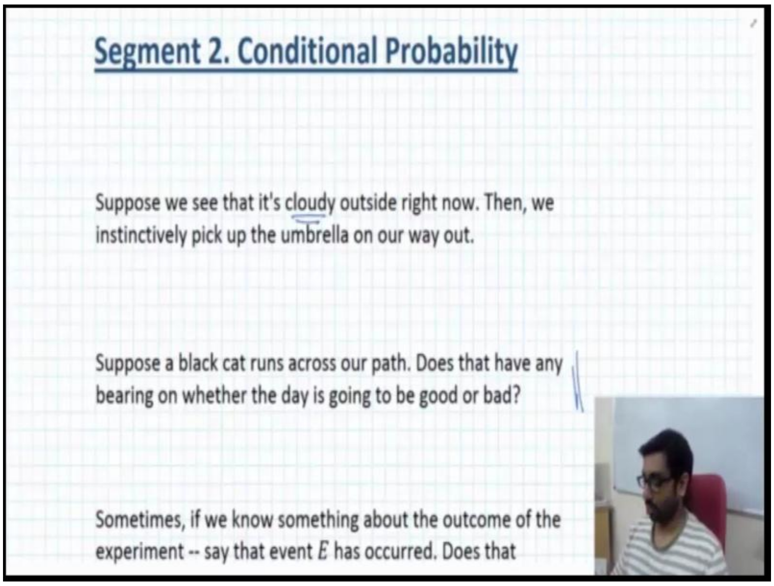
Algorithms for Big Data
Prof. John Ebenezer Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture – 02
Conditional Probability

(Refer Slide Time: 00:15)

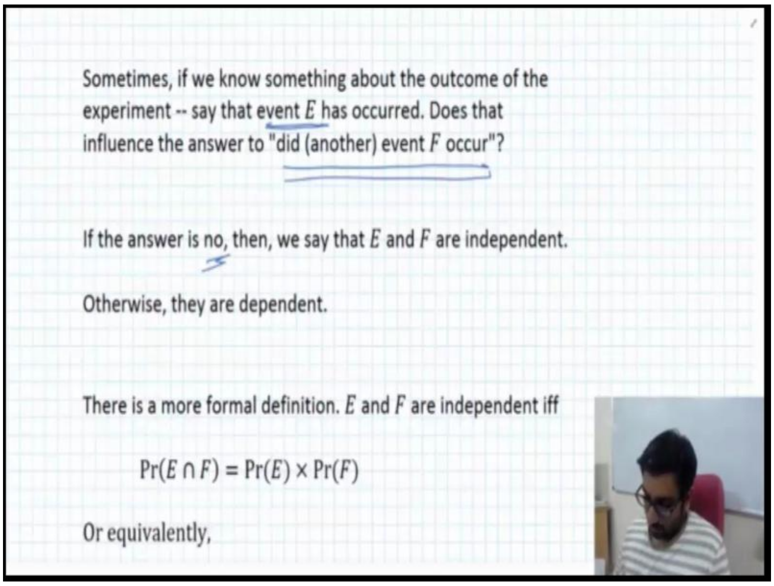


Hello everybody. So, we are in the second segment of the first lecture. The lecture is on probability theory and in this segment we will be focusing on the topic of conditional probability and let us motivate this topic by looking at a couple of scenarios.

Now, first scenario - suppose you look out of the window and it is a bit cloudy then what would you do? You would want to pick up the umbrella on your way out because if it is cloudy now, then it is likely to be rainy in a little while. If you look at another scenario: what if the black cat runs across your path? Does that have any bearing on whether the day is going to be good or bad? Now, some of you may think that it may be a bad omen, but scientifically speaking, there is no clear correlation between the two things.

So, just, these are just a couple of examples, situations, where an event occurs and it may or may not have a bearing on some other event and this notion of dependencies between events is what we are going to talk about today.

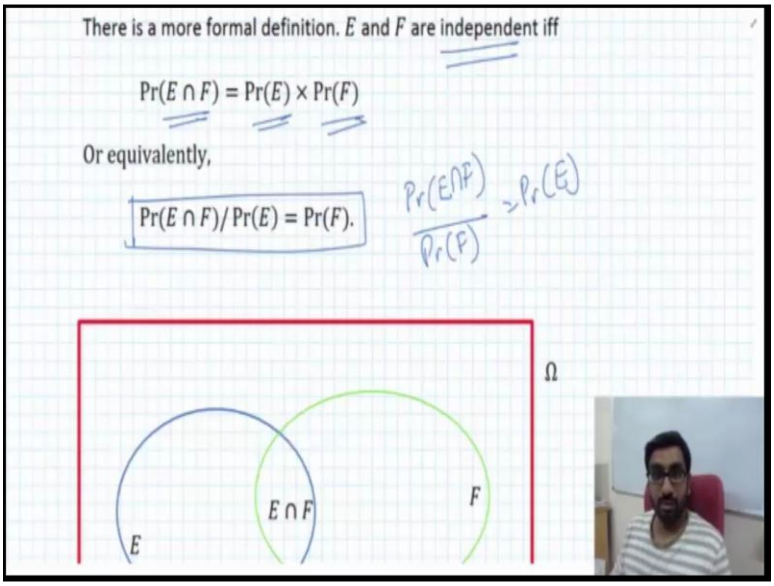
(Refer Slide Time: 01:39)



So, let us move towards making things a bit more formal. Suppose, we know something about the outcome of an experiment, in particular let us say we know that a particular event E has occurred. Now, does that influence the answer to the question, did some other event F occur? If the answer is no, then we say, that the two events are independent because now E has occurred, it has no bearing on whether F has occurred or not and so, they are set to be independent.

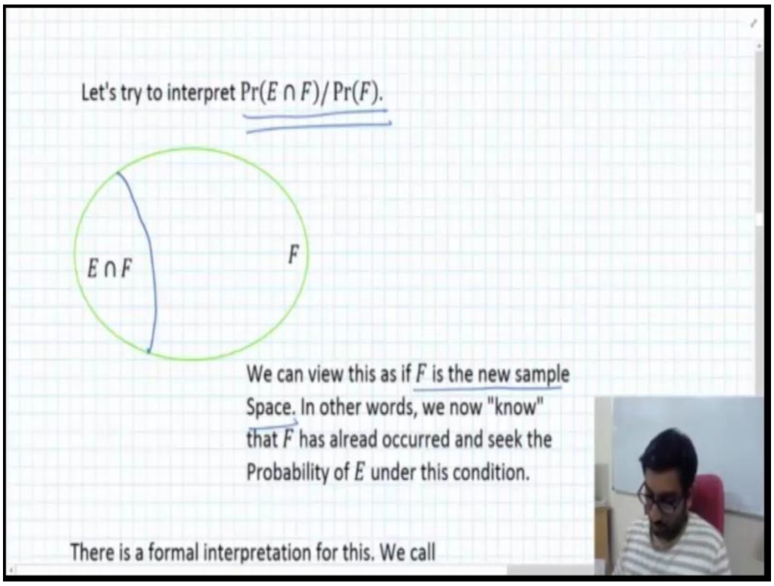
On the other hand, now suppose the event E occurs and that somehow changes the probability of F occurring or not occurring, then we say, that two events are dependent on each other. This is an intuitive definition, but there is actually a more formal definition. So, let us look at the more formal definition.

(Refer Slide Time: 02:49)



Now, we have two events, E and F and we want to know whether they are independent. Why did we that they are independent of each other? When the probability of the event E intersected with F equals the probability of E times the probability of F then, we say that two events are independent of each other. Now, the equivalent way of stating it is, of course, probability of E intersecting F divided by probability of E equals probability of F or yet another way would be, say, probability of E intersected with F divided by probability of F equal to probability of E.

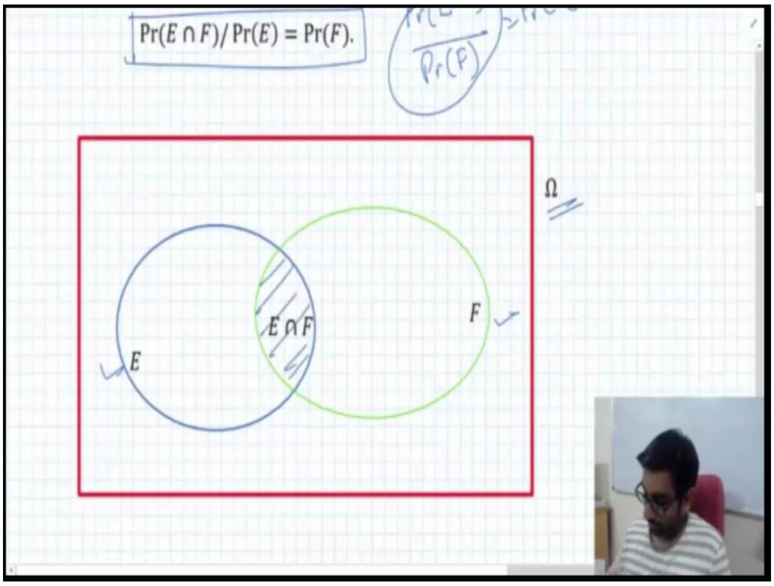
(Refer Slide Time: 04:13)



Now, if you were anything like me, it does not immediately jump out at me as to why these two definitions, intuitive definition and the formal definition are (Refer Time: 04:08). So, let us try to reconcile the two definitions. And for this is what we will interpret this, that (Refer Time: 04:21) side over here, so that is what we have over here.

Probability of E intersected with F divided by the probability of F. So, this division by the probability of F can be thought of as a sort of a normalizing the division in which what we are doing is, we are now restricting the sample space. So, this distribution by F can be viewed as if F is the new sample space. So, the probability space now looks like this. F is the entire space and E intersected with F.

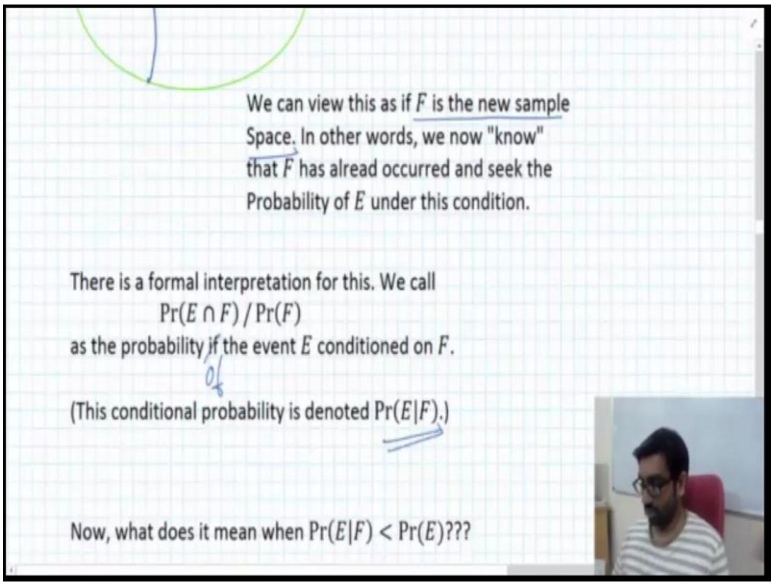
(Refer Slide Time: 05:05)



Now, previously, if you just recall the Venn diagram, over there was the whole sample space, there was the F over here, E over here and this was intersected with F over here. Now, the new sample space is just F and we have just the part E intersected with F. So, this is a new picture we get by looking at the left hand side.

And a way to think of this is, is to think if as though F has somehow already occurred and because of F has occurred, anything outside of F does not matter, that is why, the entire sample space is just F.

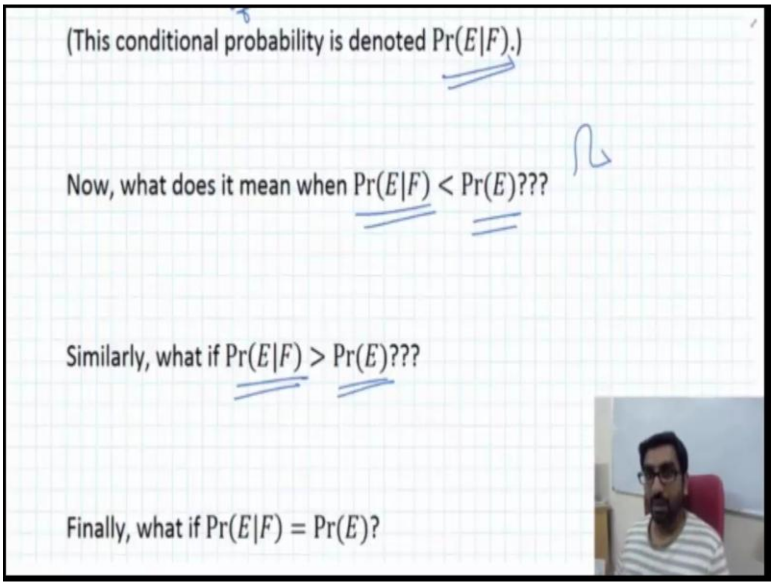
(Refer Slide Time: 06:02)



And now, we want to understand the probability of E. Now, as it turns out, there is actually a formal way to state this.

This probability of E intersected with F divided by probability of F is also called the probability of the event E conditioned on F. So, the condition is that F is already occurred. We know that F has occurred and therefore, under that condition what is the probability of event E? So, this is denoted, the probability of E conditioned on f.

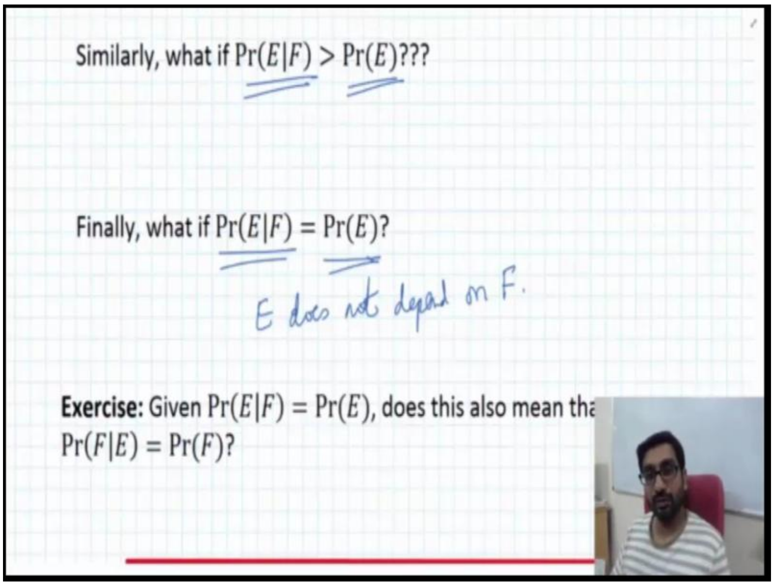
(Refer Slide Time: 06:50)



Now, that we have this insight into the left hand side of the, of the, of that one particular formula or the definition of independence, let us actually ask a few questions. Now, what does it mean when the probability of E conditioned on F is less than the probability of E? Now, here this is the probability under the restricted sample space. Now, that sample space is just F and we are asking what is the probability of E. On the right hand side is the probability of E in the full sample space of omega, okay. Now, if this probability of E given F is less, now that what, what does that mean?

If someone, if, if you know, if you know, that F has occurred, then the probability of the event E somehow is now less, is reduced. Similarly, when the probability of E given F is greater than the probability of E, somehow under this conditioning on F the probability of the event E has increased. In either of these cases, under the condition that F has occurred, the probability of event E has, has changed and this is the, this is correlating with the intuition, that we already have. So, knowing something about the event F has a bearing on the event E and so, E and F cannot be independent of each other.

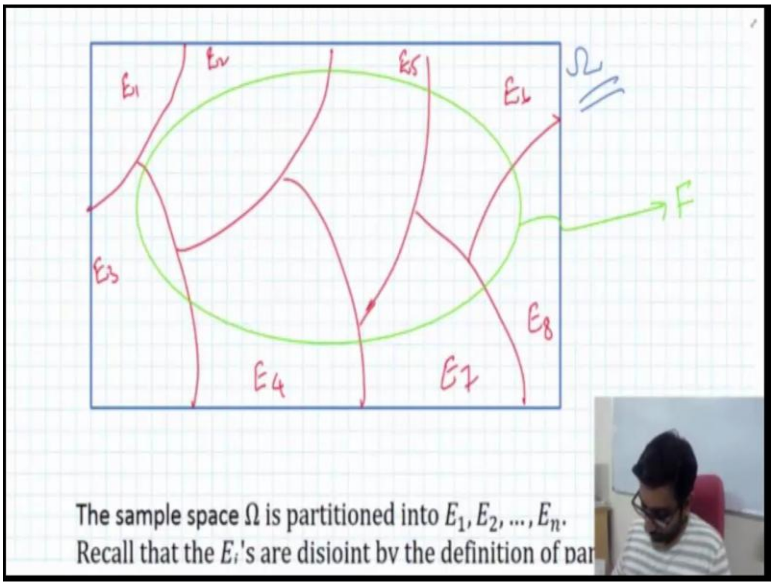
(Refer Slide Time: 08:29)



On the other hand, now let us finally, look at the case where probability of E given F is equal to the probability of E. Now, one of the conditions that F has occurred, the probability with which event E occurs or does not occur is the same, it has not changed. And therefore, now we can say that E does not depend on F.

Now, the question is, given the probability E given F is equal to the probability E, does this also mean the other way round where probability of F given E is equal to the probability of F? Now, what do you think? Try to work this out. This is an interesting question.

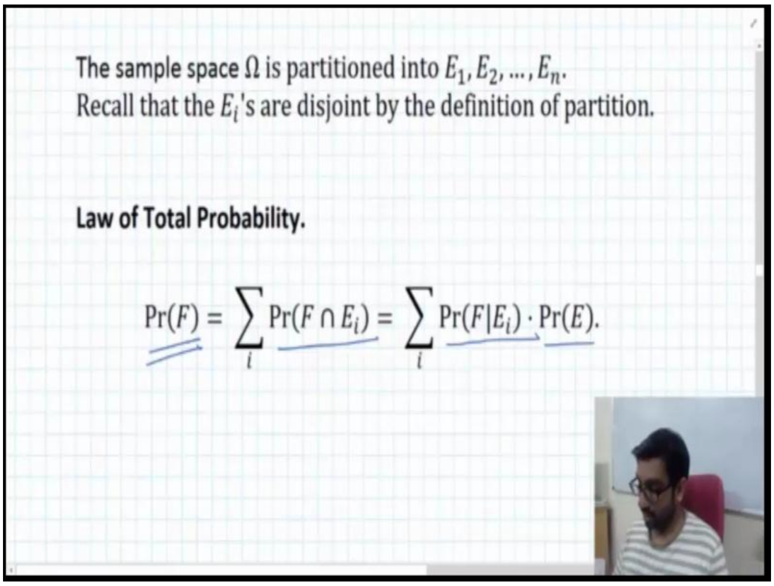
(Refer Slide Time: 09:24)



Now, let us talk about an interesting and useful law that uses conditional probability and this law is useful in the context where we are interested in understanding the probability with which a particular event F occurs, but it occurs in somewhat of a complicated sample space and we do not have quite a handle on the, on the event F, but we have a way of, let us say partitioning the sample space into lots of different events, E 1, E 2 and so on. And under these individual events, these, these disjoint events, we are able to understand something about, about F.

So, this is a picture. We have, we have the sample space the omega and it is partitioned into E 1, E 2 and so on, to E 8 in this picture and we have the event F and the event F intersects with some of the E i's and not with all of them. And so what is this? So, how do we now get an understanding of the probability of event F under this complicated scenario? Now, the assumption here is, that under each event E i, we have some understanding of the probability of F. So, let us, let us see how this plays out.

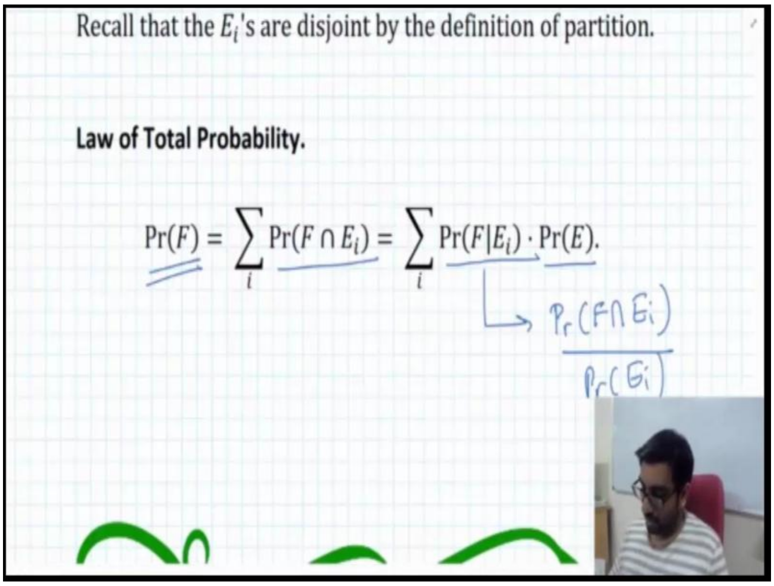
(Refer Slide Time: 11:00)



So, we are of, as, as I mentioned, we are interested in the probability of the event F and this probability, I will, again going back to the very definition of probability function should be the summation of the probabilities of F intersected with each of the E i's, why, because the E i's are all disjoint and therefore, F intersected with E i's. They are all, all going to be disjoint and together that, those events are going to, if you, if you unite all of those events, F intersected E i, they are all going to equal F. And so, the probabilities summed up should be the probability of event F and this comes back, comes from the very definition of the probability function; great.

And now, we simply notice, that the probability of F intersected with an E i can be written as the probability of F given E i times, the probability of E. Why is that? Well, you may recall, that the formula for probability of E given F here is probability of E given F divided by the probability of F. We are just going to apply that formula here.

(Refer Slide Time: 12:40)



And so, this one is probability of F intersected with E i divided by probability of E i and this just moving the probability of E i over to the other side, we get what we need. And so, now where does this law of total probability as this law is called, where would it be useful?

Now, as I have pointed out little while ago, that we, it is useful whenever we have a handle on the probability of F given a particular event E i and once we have such an understanding of F for each of the parts E i's then we can put them out together and get an understanding of the total probability of the event F.