Introduction to Stochastic Processes

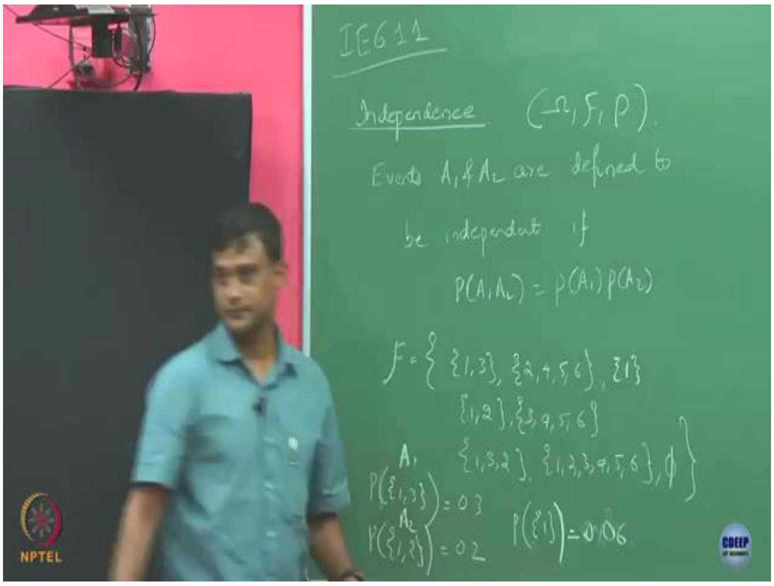
Professor. Manjesh Hanawal
Industrial Engineering and Operations Research,
Indian Institute of Technology Bombay.

Lecture 03

Independence of events and Conditional Probability

So, we are going to discuss today two important properties called Independence of a Set of Events, then we also going to talk conditional probabilities and then we will define what is a random variable today. So, if you recall in the last class we just begin with what is a probability space, we defined what we mean by sigma algebra, what we mean by a probability space. So, that is the basic building block for us and we keeping on building on this aspects.

(Refer Slide Time: 01:07)



This, notion of independence of two events is important in our study of probability and this notion of independence you have go purely by definition, the definition that I am going to give you. We say that, so now on words whenever I am going to define this thing all this things, I am going to assume that these things are define under some underline probability space which has its own simple space, F, even space and then associated probability function. So, let us say this is given.

Now, I am going to say, let us take two events, coming from my event space, I am going to say that they are going to independent if the probability such that, the probability of a joint event split as product of the probability of individual events. So, this is just definition, so independence has nothing to do here but these A1, A2 are disjoint or anything of that sort, it is just the property of my probability function.

For example, let us take our dice example; in our dice example let us construct one event space, which will have these events. Let say I have such an event space, right now I am not worried about, it is a sigma algebra matrix already sigma algebra, but I am not worried about it.

And on this let us define my probabilities like this, my probability I have defined, the probability of the event 1, 3 to be let us say I have defined this to be happened to be 0.3 and I have define 2.2 and I had defined probability of 1 to be 0.6. And let me also include this event 1 here.

Now, according to this, through definition of my probability, are my events 1, 3 and 1, 2 independent? So, let us take this is to be A1 and let us take this to be event A2. I have defined these probabilities to be, what is the event A1, A2? When I say A1, A2 it means interception of A1 and A2, what is A1, A2? And what is the probability of that event? 0.6. Now, is this 0.6 is going to be the product of this two?

Student: sir, no.

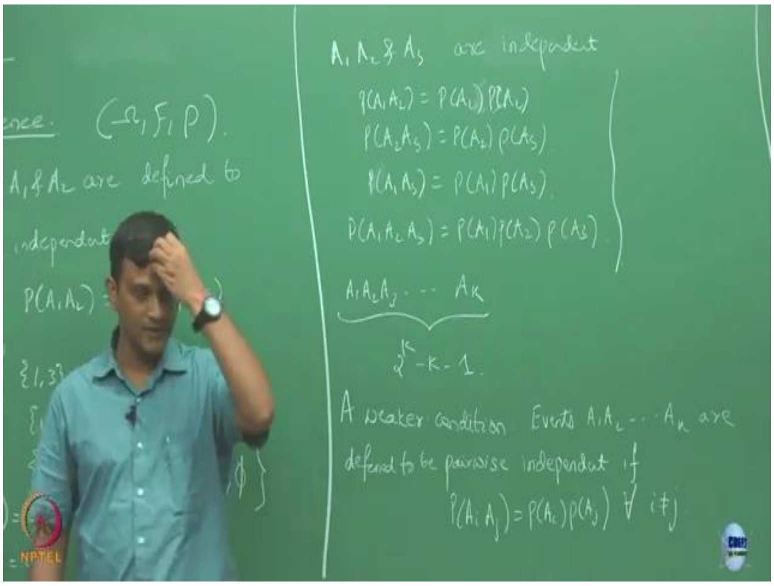
Professor: No right. So, in this, you then say, are you going to say A1 and A2 are independent?

Student: No.

Professor: Suppose if I redefine this to be 0 point 06, then they are independent under this probability.

So, as you see this, it has nothing to do with about any commonality that are occurring on this two events. It we are, according to this definition we will simply go that, as long as the probability of this joint event splits as the probability of individual the events, we are just going to called that as independent. Now, this is the notion for independent of two events.

(Refer Slide Time: 06:29)



Now, can we talk about independent of three events? How about, so when you are going to say A1, A2 and A3 are independent, they are going to say, they are going to be independent. Now, if I take P1, P2 this should satisfy, if all this conditions holds, then we are going to say that the events A1, A2 and A3 are independent.

So, suppose now I say that, if I now I want to extend it to A1, A2, A3 all the way to, there are Ak, these many sets are there. What you think the independence of this many event means? Here I took pairwise and also took, so what I did basically here? I took all possible subset. So, then when you have A1, A2, A3 we are going to look for all possible subsets. But in the all possible subset you do not need to check for individual elements, that P of A1, P of A2 you do not need to check, because all subsets also includes individual elements.

Now, if how many such conditions over all, how many conditions will be there? If I had to check the independences of this many events, how many condition you think you need to check?

Student: (())(08:50)

Professor: 2 to the power k minus k, why 2 to the power k first? So, this means basically total number of subset, and why k you have to remove? Single individuals I have to remove and this 2 to the power k also contains what? My set, I do not need to check anything for null set so have. So, if you want to check independent of k events you need to be checking this many conditions.

Student: Sir, k are the individual single turn sets sir.

Professor: Yes, there are k single turn sets.

And I am asking if you had to check independences of this many event, then you are going do such things for all possible subsets and the actual number of subsets you are to need check is this many, so you will be ending up checking this many conditions.

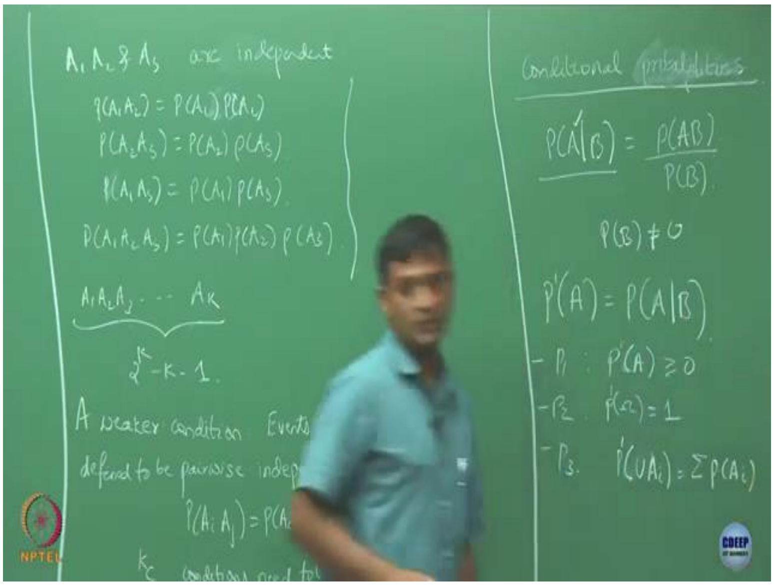
So, as you see checking for independence of k sets already requires large number of conditions to be checked, that is like it is growing exponentially like to the power k. So, often looking at independence off, it is kind of joint independent, instead of looking at such kind of independence we will be interested in looking at pairwise independence.

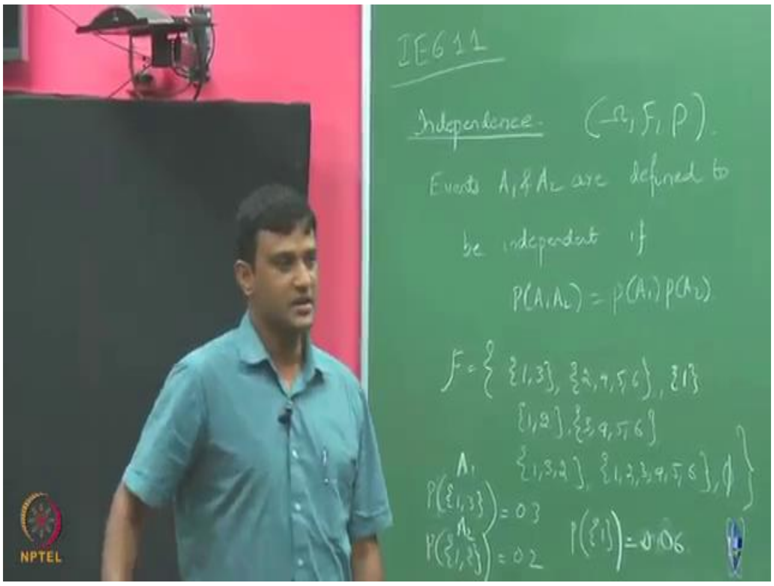
Which we call a weaker condition. So, when I want to ask check you check only for the pairwise independence of this k events, where only take when to take two at a time and check whether they are going to independent. So, to check pairwise independence how many conditions you need to check?

Student: (())(11:35)

Professor: k c to

(Refer Slide Time: 11:56)





So, now we are going to define the notion of conditional independence. Before that, we will define conditional probabilities. So, when you are going to model many thing, anything you will of often will be face with a situation that, what happens if this event has happened, suppose let us say there are many things that are going to effect that are random things, but you observed, that some event already happent, now conditioned upon this event now you want to analyse the system.

For example let us say, you have to dice, you are going to throw them, you are going to throw first one an let us say after throwing the first you are going to throw the second one, this is one experiment for me, throwing that dices, one after another. So, what are the sample space for this? How many points are there? There are 36 sample points are there in this sample space.

Now, let say as soon as I throw the first dice, I will already observe what is the outcome of this dice. Now based this one, there is still randomness in this, because I have not rolled my second dices. Now, but my randomness has significantly reduce now because I have made some observation already about the first one.

Suppose if you know the first one, what will your sample space reduces to, it is just going to reduce to 6. Now you have to just look at the randomness that is going to appear out of that can possibly come from this 6 outcome. So, you see in that in this case after observing something you want to condition of thing and look at a new induce sample space, after making this observation.

And now that will lead us to something called conditional probabilities. So, we are going to say that probability of A given B. What is the meaning of this? Probability that event A happens given that B has already happened. And now we are going to define this as probability of AB.

So, now in the case of, suppose if I tell you, what is the probability that the some of the outcome of a dices is 10 given that the outcome of the first dice is 4. How you are going to compute this? So, I already told you the outcome of the first dice says 4. Now you are only in going to look at the possible outcomes which are 4 1, 4 2, 4 3, 4 4, 4 5, 4 6.

Now, in, now you that, my sample space has reduce conditioned on this. Now, in this, now you will be looking at event of interest that is the outcome of the dice is 10, in what way this can be possible?

Student: (())(15:41)

Professor: This is possible only if 4 after I observe 4 I observe 6, that means out of the restricted set I will be now looking at the occurrence of this pair 4 6.

So, this is exactly we are going to say. Now, that the event B has happen, related to this event B, what is the frequency of event of my interest? So, let us say, I am interested in some event A, after event B has happened, if this A has be of any interest this event A has to be a part of event B, because if that event is not part of B, this is never be any event.

For example when I say, outcome the some of the two dices is stand that is the pair 4 6, this is one of the possible candidates in 6 possible outcomes I had. In that what is that event of interest that is common to both A and B and what is its relative frequency in the new reduced sample space I have. So, this is how I have we are going to redefine our conditional probability, given that my probability has given that some event has been told to happen, that has happen that is going to happen with probability PB.

That means when you observe that particular event, that as it wants probability and that would have happen with PB, within in this probability what is the relative frequency of that I am going to observe event of my interest and that is how we are going to define this conditional probability and this conditional probability in discussion.

Now, naturally I have to see this conditional probability you want to define for something, the event on which you are conditioning that happens with some positive probability. For example, in the dice case you are not going to condition on the case flare, the outcome of the first dice is 0, that is not a possibility for you. So, I am going to assume that for that reason we are going to going define probabilities, condition from the events for which the probability of the events is non zero.

Now, as you see this, once I fix my B, now I can change my event A, whatever event A, I can ask, for example in the coin toss problem let us say first the B event B I said the outcome of the first event is 4. After this I could ask the question, the sum of the outcome of the dice is 6, some of the outcome of the dice is 7, some of the outcome of the dice is may be 12, I can ask all this questions and all this corresponds to different events A.

Now, each of these events will have their own associated probability, condition on this B. Now, you can ask the question, P was a probability space which I have started with, which has, which need to satisfy certain set of axioms, we said P1, P2, P3 are the condition that need to satisfy.

Now, If I am going to treat this guy as a new probability that is you fix this B underlying event B and then ask for the probability of events A. Now, let us called that, now this is another probability, (())(20:03) on different different event, all these events are also coming from my script F, is this again a probability function? If I need to check this, is probability function what I need to verify?

Student: (())(20:19)

Professor: So, I need to basically verify type three properties P1, P2.

What was the P1 property it says basically non negativity. So, is my P of A is going to be greater or equals to 0? Yes, right this is by definition. So, if you had define it like this P is always going to give positive values, this is going naturally and what are my second probability? Can you somebody tell me precisely what will be is the second one?

Student: probability of (())(21:00)

Professor: What you mean? Is this property holds? Why is that? So, just plug in for A equals to omega here. What is going to be, if A is omega, what is omega intersection B? B. So probability of B by probability of B you are going get it as 1. What is the third one? So, we want that, if let us take a sequence A1, A2, we want this to be what, we want this to hold for countably many A I's. is this true?

So, just verify this condition as well. So, because of this, we will going to say that this conditional probability. So that is why, even though we just define this in this function this conditional events based on this, but you are we still called it conditional probability, because it is already satisfying this properties of probability function. So, PI is probability function.

(Refer Slide Time: 22:59)

Now, from the independence property like I had just said, I had given you the definition that, what I mean by A1 and A2 are independent function. Suppose I say I am going to denote independence by A1. What is this mean? What I under that probability function, when I say A1 is independent of A2 that means probability of A1, A2 is going to be probability of is equals to probability of A1 times probability of A2.

I am just going to use this short hand notation for that. Now, suppose A1 and A2 are independent, does this imply that A1 compliment is also independent of A2, just apply your definition of probability and see whether it works. So, what I want to show, If I want to show this is true, I need to show that P1, A1 compliment of A2 is equal to P1, P1 compliment and P of A2. So, this is what I need to show.

And what I know? From this I know that, I know that P of A1 of A2 is equals to, if I assume this is true that is A1 and A2 are independent, if this is going to be true. Let us try to see why is true. So, now let us say I have this event A1 here and event A2 here, let us say they have some overlap that is fine. So, now what is event A1 compliment A2 what has the region that it includes?

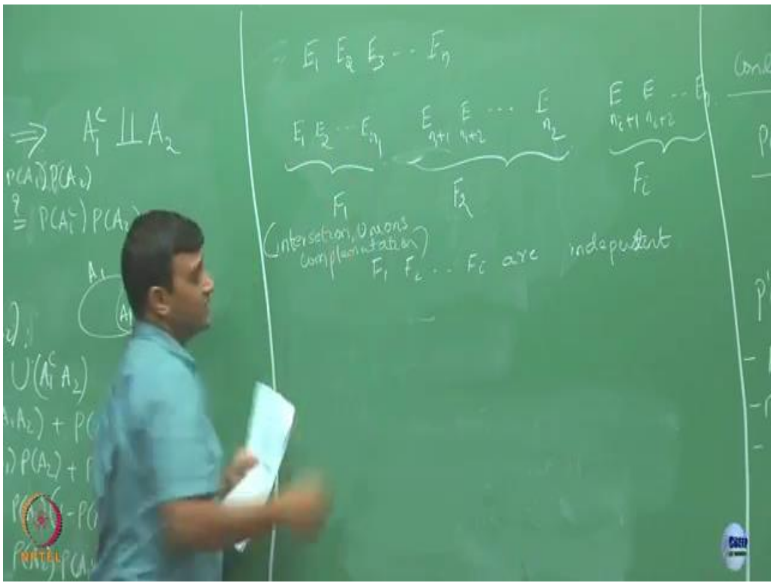
So, it includes this region. This region how can I write it as union of two disjoint regions? So, can I write it as, what is this region? This is A1 intersection A2 this part. Now what I want is A1 compliment A2. So, you can check this, I can write it as this A2 as simply this guy, A1 intersection A2 and then union of whatever I am interested in, is this correct? I am just written and are this two of disjoint?

If they are disjoint, I can just apply the probability of A2 to be the some of the probabilities of these two terms. So, then, now what I know, now I want to I have been told A1 and A2 are independent, if A1 and A2 are independent I am going to apply that condition here, what I am going to get? Why 0?

P of A1 into P of A2 whatever this part is. Now, if you simplify, take this quantity on the other side, what you are going to get? P of A1 compliment of A2, this is going to be what P of A2 1 minus P of A1. This is correct? And what is 1 minus P of A1? It is simply going to be P1 of compliment.

So, you can see that like A1 A2 are independent you can simple derive many such properties like that. In this case I have derived one simple properties like A1 compliment and A2 are also independent.

(Refer Slide Time: 28:16)



So, just to you can keep on extending this analogy, just to let see, if you guys can understand this, suppose I say that E1 E2 and E1 these are n number of events, there are totally how many events? There are n number of event. Suppose if I do a partition of them by taking E1. I am going to take E1 E2 all the way up to first n 1 numbers. Let, us say n is equal to 50, I will take E1 E2 up to n1 I am going to set as 10, first 10 sets and then I am going to take it as En1 plus 1, En1 plus 2, I am going to take all the way up to En2.

Let, us say n1 stand this is 11, 12, n2 is may be, n2 is 30 till that. And like that I will I will make a whatever the remaining ones. Let us say similarly, I have going to get after and I plus 1, E n plus 2 all the way up to En. So, I basically group them, like this one consist of n1 elements, this one consist of certain number of elements n2 minus n1. And this will consist of another set of elements, this total number of n.

Now, suppose from n1 numbers, you take whatever union complementation of this sets like this, for example like here when I have two, I had taken A1 compliment and A2, like that you take arbitrary things like may be you can take E1 compliment intersection E2 intersection E3 compliment like this any combination.

And whatever the resulting function, let us called F1 after doing that operation let us call this another second operation and let us called this as i th operation. Now, you can argue that, if E1 E2 up to E1 are independent then F1 F2 Fi are independent. You got the sense what I am trying to say, if you take any set you do arbitrary operations on them complimenting intersection and then.

Let, us call that F1 and take another set and do, if you are going to look at this sets now. This is one set now, if you are look at this sets now they are, themselves are independent. So, when I say I got this, this are any Boolean functions like you can do intersections, unions and complementation of your this sets and get a function F1. Similarly you do this and get F2 and like that. It is up to you how you are going to define how many sets you want to take to get F1, how many sets you want to take F2.

As long as this guy's event is, are F2 this F1, F2 all the way to Fi will also be independent. So, as you see independence is actual very strong property, because to satisfy this independence you, for k elements you already need to (())(32:32) 2 to the power k minus k minus 1 number of conditions. So, because of that I would like that n you have set of things and if you do operations among them you will end up with sets which are further dependent of each other.