Engineering Statistics

Professor Manjesh Hanawal

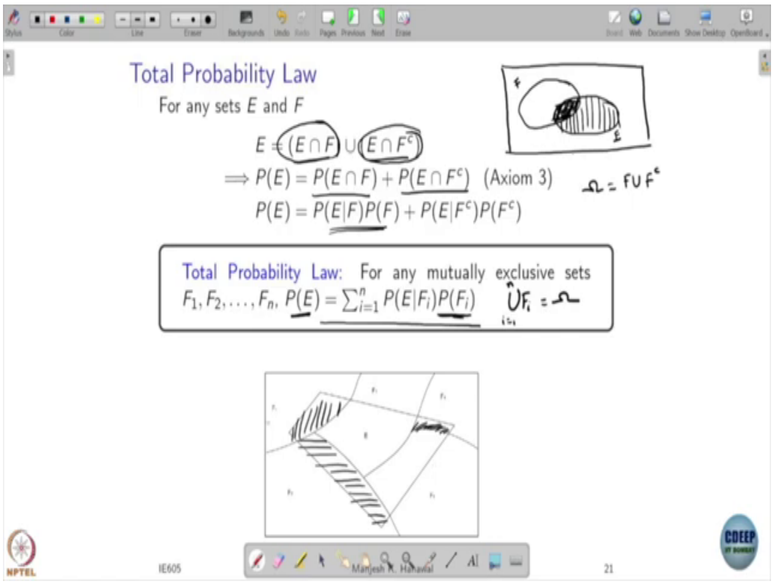
Department Industrial Engineering and Operations Research

Indian Institute of Technology, Bombay

Week 01- Lecture 05

Total Probability law and Baye's theorem - II

(Refer Slide Time: 0:18)



Now, next I want to talk about one important law in probability called total probability law. Just to understand total probability law let us take two events and let us say there is one set F and another set E and this is F. What I will do is I know that omega can be written as $F \cup F^c$, F and F^c is the entire thing.

So, now what I am going to do is? What portion of E lies in F that is $E \cap F$ and what portion of E lies in F^c that is E intersection F^c and that is exactly E. All of you agree? Now, if you look into that this portion and this portion are they mutually exclusive? So, what I have basically done is in E I have looked into that portion that overlaps with E and the remaining part outside and this portion, so this portion of this, sorry this portion of this lies with F and the vertical portion lies with F^c and we know that this vertical portion this horizontal portion they are mutually exclusive, this region.

Now, if it is the case now I can if I want to compute the probability of these things I can apply the third axiom of probability which said that if two events are mutually exclusive the probability they are union is nothing but some of their probabilities, so I will do that, probability of E is probability of this part and probability of this part.

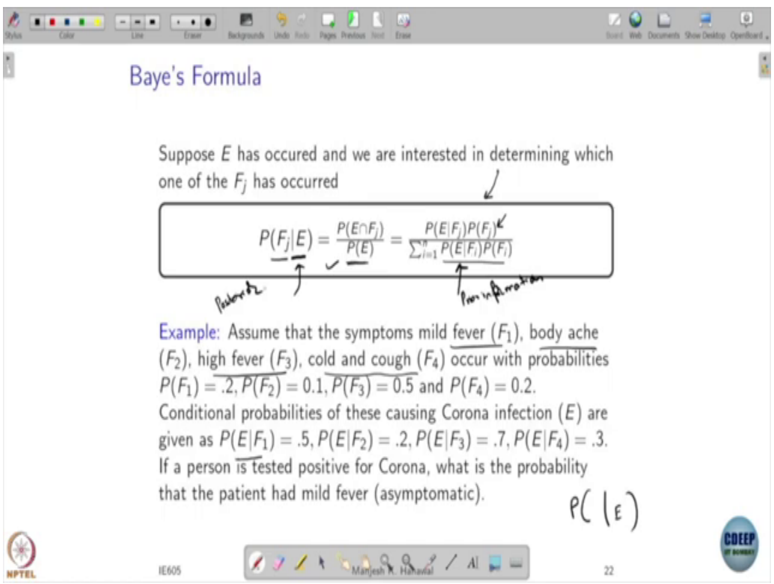
And now I will apply the conditional probability on this by definition probability of E intersection F is nothing but probability of E given F into F and probability of F given F complement into P of F^c, this I have done it for two things, two events but assuming that my like this I have assumed like here the conditioning happening on F and F^c, but I can go and generalize this.

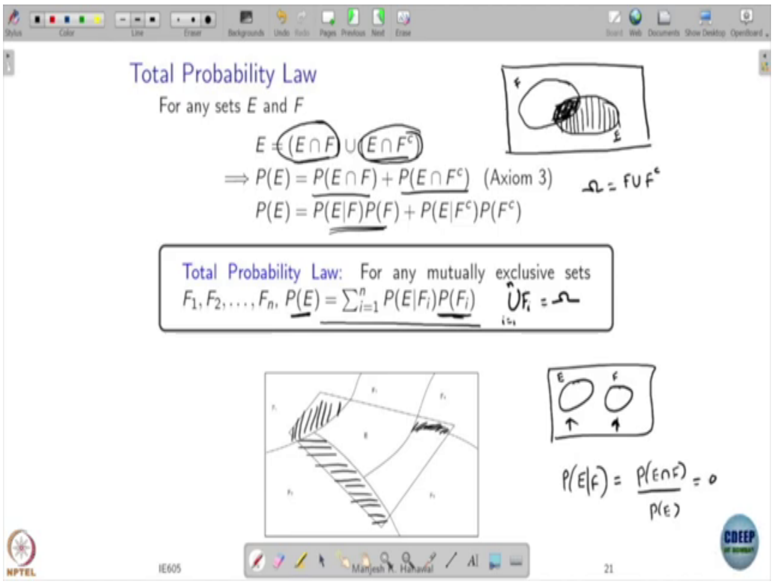
Suppose let us say I have F1, F2 and FN which are mutually exclusive and one notion I have also missed is like and these are also a partition, like union of F is equal to omega, i equals to 1 to n. I have this partition then for any E you will be able to write it like this and this is called total probability law.

What total probability law is saying you is you can compute the unconditional probability of E based on the conditional probability of E on this F i's which forms a partition after multiplying the probability of those partitions. So, pictorially that is what it is saying I mean what I showed it for two simply F here I have now taken this region F1, F2, F3 is this, F4 and F i and now this there is some event in this grey shaded area E and I want to if I want to compute this probability of E, E could happen because of this portion that is falling inside F1 or E could have happened because of this portion that is falling in F4 or like maybe this portion which is falling with an F2.

So, I am just taking all of them and adding them here. If you recall I said that event E happens when some one of the possibilities in the event E happens and it can happen whatever that happen can fall in any of the partitions.

(Refer Slide Time: 4:58)





Now, using this we have now one fundamental formula in probability called baye's formula. And now let us say suppose E has happened and now I want to, now E has happened now I want to let us say now I know that let that event E has happened now I am interested whether E happened because of the things that are happening in the F4 region or F1 region or F3 region like that I can ask the question and that is what we are trying to now see.

If event E has happened what is the probability that it is coming from the Fj part now there we can again go and apply our conditional probability formula and conditional probability again P($E \cap F_j$) again numerator is again I am applying the conditional probability now I am basically conditioning on F j and denominator P(E) is coming from the total probability. If you notice that on the right-hand side everything is like conditioned on F j, E is conditioned on F j and now we are (condi), on the left-hand side it is conditioned on E.

Student: (())(6:24) if there is no intersection…

Professor: So, then what is that if and like this, now what is the probability let us say this is E and this is F and now you want to compute probability of F. Now, what is this? Just go and apply yours E intersection F divided by P(F). What is P intersection F in this case? Null set so it is going to be 0. Because now if I tell you that F has happened can E happen? No, because E is not overwrapping with F at all. So, if I know that F has happened there is no way E can happen so it probability is 0, if they are disjoint if you condition on 1 you are going to get a 0 value.

And now, I mean here is one simple example I will just read and give it to you for verification. Suppose let us say I mean this example came when we are doing all the things on online when we are all fighting the corona thing. Suppose let us say there are three symptoms that are possible, one is like you will get a mild fever and another is like you get a body ache and you will get a high fever, these are the and like cold and cough.

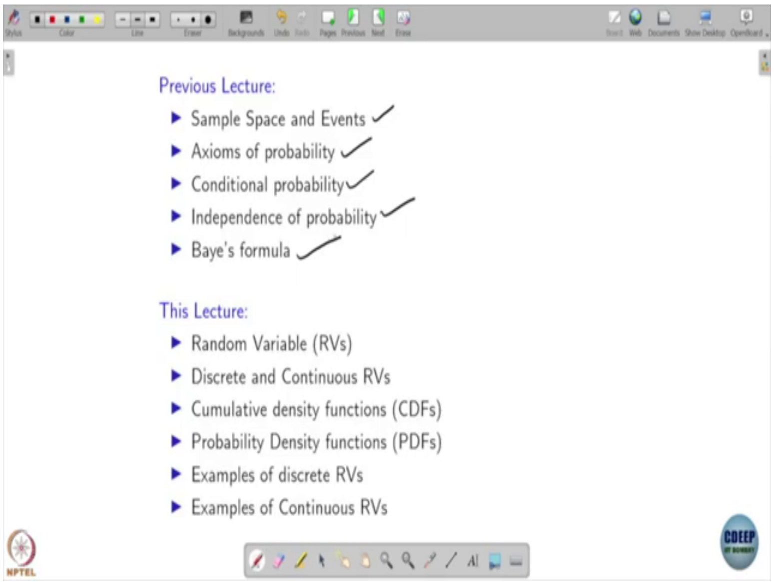
And there are four possible symptoms mild fever, body ache, high fever and cold and cough. And there is some probability of all of them happening and we were seeing all these symptoms when during the peak of the pandemic and all of them we wanted to see that like and we wanted to see if somebody is really positive it could be because of which symptoms or what kind of symptoms you will have.

Now, suppose each of this happening is like a probability of F1 is 0.2, probability of F2 is 0.1 like that. Now, you want to see that if somebody is having a mild fever what is the probability that you would, you will have an infection. So, to do this, okay, now let us say we know that like if somebody has a mild fever then probability that will be infected is 0.5, somebody has a mild, sorry what is a body ache then the probability that he is going to have infected is 0.2 like this.

And now let us say we want to now con, ask the reverse question, somebody's event E is like he is already tested positive. What is the probability that if somebody’s test is positive because of he having one of the symptoms and now you can see that you can go and apply this Baye’s formula there.

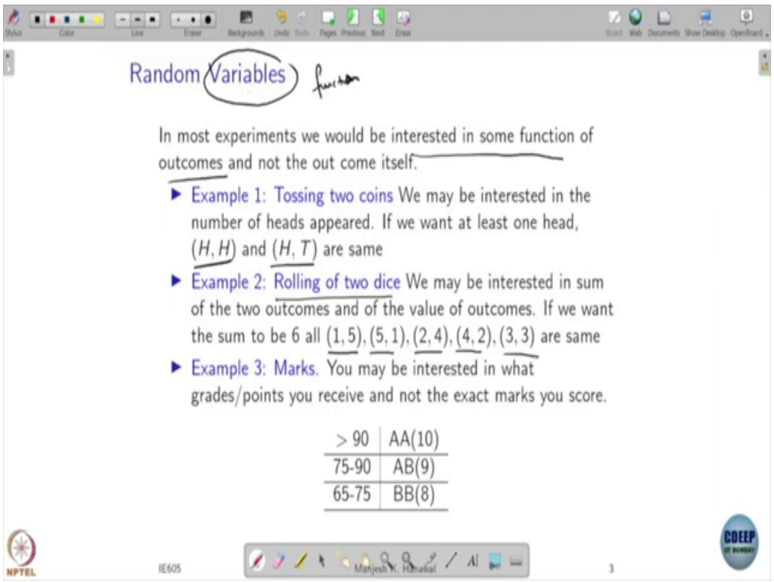
And the Baye’s formula has many many applications. So, usually what we call this as like a kind of prior information. Like often we say that we know something and now after the actual event let us say here (actu), event E is the actual event that has happened. So, we know that event E happening if something has happened like F i's we have some information, now that event F itself has happened we want to compute what is the probability that it came from event F i and this is sometimes called like a posterior. So, now you are trying to find out posterior probability using our prior information. So, fine. Let me see if any of you have any questions, if not let me move to the next one.

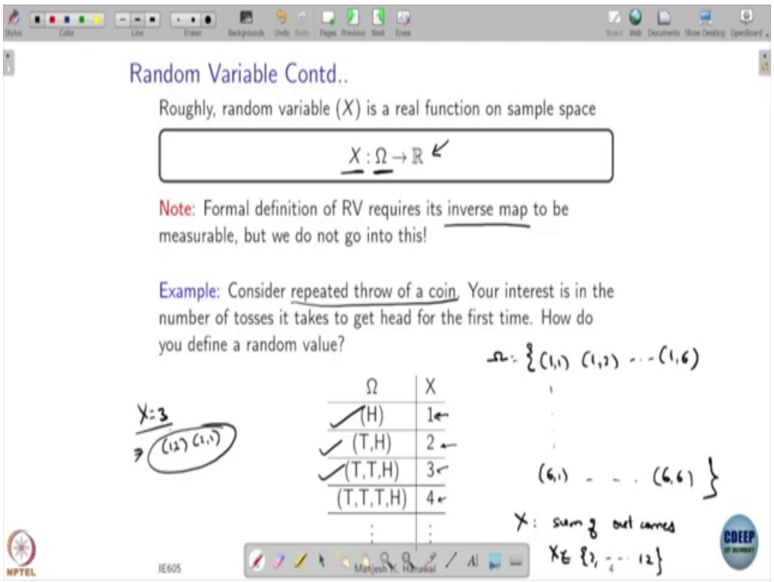
(Refer Slide Time: 10:38)



So, we just went through the following things, sample events, axioms of probability, conditional probability, independence of probability and Baye's formulas. Now, what we will talk about is random variables, discrete and continuous random variables and some of the functions associated with them, called CDFs and PDFs.

(Refer Slide Time: 11:00)





And where is the random variable coming into picture? So, in most experiments we would be interested in some function of the outcomes and not the outcome itself. Why? Because when you are going to throw a coin, what you are going to see is heads and tails but you are may not be interested in heads or tails. What you may be interested in number of the times head comes or number of the time tail comes or number of trials you have to throw till head comes, you will be interested in such questions that are related to numbers.

Now, that is where random variable comes into picture. We want to map the outcomes to numbers, instead of like outcome could be like head tails and all I am, I want to get rid of that and everything I want to convert to numbers so that I can do some mathematical operations or numbers I can crunch numbers but not heads, tails and some characters.

So, let us look into some examples like tossing two coins, we may be interested in the number of heads appeared. And in this case if, in this case if I am interested in heads number of heads and I am interested in at least one head when you throw two times both of them are fine, in both the like here head had come, here first head had come and tail has come later in the second toss. But in both of those one head has happened so they are fine with me.

So, that is what like actual outcome is not of consequence to me, actual outcome but what was my interest is whether one head has happened that number one is of importance to me. Now, if you look into the rolling of a dice, you may be interested in outcome being 6, outcome being 6 could happen because first one showed phase 1 and the second one showed 5 or reverse or it could be like a 2, 4 or again that get reverse and 3, 3 any of them will lead you to number 6.

But I am not interested in specific outcome. What I am interested in my quantity of interest whether the sum is 6 and for me in that case all of them are equivalent or same. And other example could be like suppose let us say you are grading in the class like and the policy we have uses these grades like we will not assign you the exact value number you score in the class, what we will put you in the brackets of AA, AB, BB like that.

And A could be like anyone like who got 90 could be AA, and anybody who got let us say 75 to 90 bracket is like AB like in that case what matters to us is like only this ultimately these numbers 10, 9, 8 like not the exact number between 0 to 100 you got. So, we are trying, this is basically random variables like a function which is mapping the outcomes so in the class like outcome is the values you score but we are now mapping it to some scaled value between 0 to 10 instead of taking 0 to 100, that is where we are now getting into the random variables.

And it is like random variable is actually kind of misnomer for some reason this variable has stuck it is actually function, random function like why is that actually a random variable X is a function from your sample space to real number, it is going to assign a real number to every possible outcome. So, this is a simple definition, so you can treat any random variable is something which assigns values to my outcomes and this definition holds when the things are discrete but when the things are continuous one has to be little more formal and we have to need to worry about the inverse maps are and all but we will not get into that.

For our purposes we will just take it like maybe this has been discussed in IE 621 like is the formal definition of the random variable discussed there about the inverse maps and all. So, fine, for time being just take that random variable is something which is a map from sample space to real numbers.

Now, let us go back to our example and see what is the random variable we can define. So, let us take a coin and I throw it repeatedly and I am interested and I will stop when the first head comes, just think that you want to try for something and you keep on attempting till you succeed and when you succeed you stop.

Now, till you stop what are the, what all the things can happen or what are the possible outcomes in this experiment, you may succeed in one attempt then that case head has come first, you may succeed in the second attempt in which first is failure second one or you may succeed in the third after two failures like that it can go on.

But here if you look this is actually my sample space like H, T, H, T, T, H, T, T, T, H, T, T, T, T, H like that it can go on, but I am not interested in H, T this sequence. What matters to me is how many attempts I took to succeed, one attempt, two attempts, three attempts and H. And here my random function omega what it is doing is for the outcome H it is assigning number 1, for T, H it is assigning number 2 and for T, T, H it is assigning number 3 like this.

So, any question on random variable here so far? Okay, let me ask this question so in this fair of two-coin, dice where we have 1, 1, 1, 2, up to 1, 6 all the way from 6, 1, up to 6, 6. I am going to define my random variable X on this sample X as sum of outcomes, I am not looking into what each phase is showing, I am interested in their sum. So, what is the possible values X can take? It can take all the way from X to 2 to 12, so X can take value from 2 to 12 like this. And now if I ask what is the probability that X equals to 3 how, can you compute what is the probability that my random variable can take value 3?

Student: 1 by 8.

Professor: So, 3 can happen because of 1, 1, sorry 1, 2, 2, 1, only these are the possible outcomes that would have given me X equals to 3 and I can then based on this I can go and ask the probability X is 3 and I can compute from these possible outcomes.