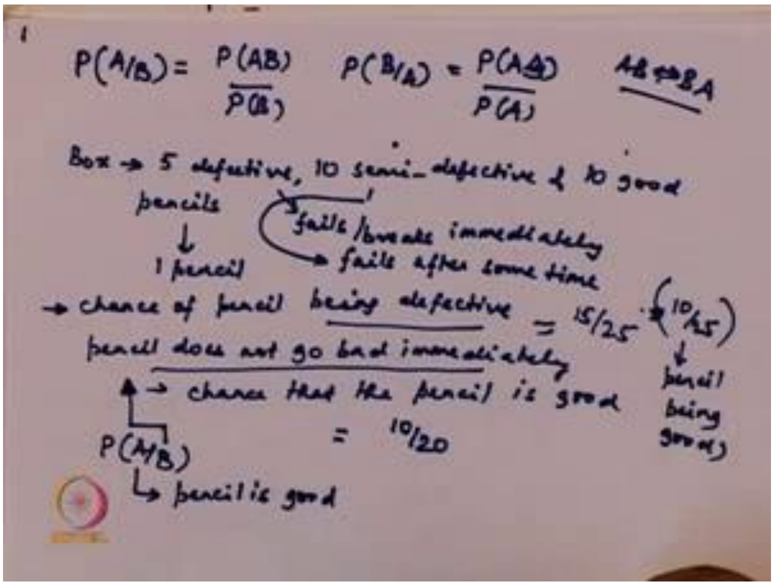
Introduction to Biostatistics
Prof. Shamik Sen
Department of Bioscience and Bioengineering
Indian Institute of Technology, Bombay

Lecture - 18
Conditional probability and Random variables

Hello and welcome to today's lecture. We will continue from where we left off in previous lecture, which was discussing conditional probability; one of the most important aspects of probability. So, how do we define conditional probability we can do a brief recap.

(Refer Slide Time: 00:34)



So, we define probability of A given B and this can be calculated as probability of A B by probability of B similarly you can write probability of B given A as probability of A B or B A by probability of a, because probability of A B B A is the same thing. So, let us consider a simple example. So, imagine you have a box, you have a box which has 5 defective, 10 semi defective, and 10 good you know pencils. So, by defective I mean that when you start using the pencil, it will keep breaking. By semi defective I mean that, when you start using the pencil it will work for some time and then fail, and good means pencils which do not break so frequently.

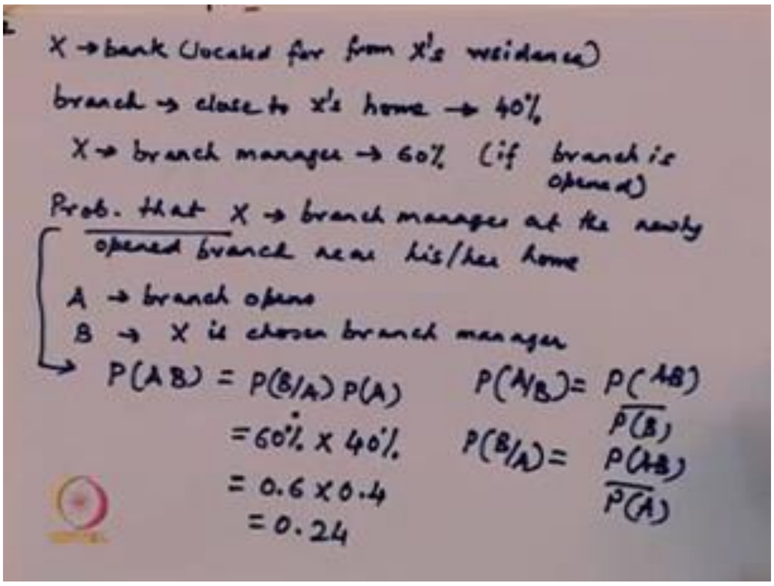
Now, let say you have taken up a pencil from this box you took a pencil one pencil, and you asked. So, what is the chance of the pencil being defective? So, chance of pencil being defective, or now being defective would include both of them, because eventually this will also fail. So, this one fails or breaks immediately, this one fails after some time, and this one these ones are good. So, chance of the pencil being defective, is essentially you want to count the total number of defective pieces. In that case it is simply 15 by 25. So, you have 15 defective and 10 good. So, event in I want to modify this particular question I say that, I take one this pencil, and it is not immediately going bad.

So, the pencil does not go bad immediately, in this case what is the chance that the pencil is good. So, I know. So, this is an example of conditional probability. I know if I have this statement, pencil does not go bad immediately, it means that I am I have either chosen a semi defective piece or a good piece. So, under this case what is the probability that the pencil is good? So, in this case it was simply 10 by 25. So, this is chance of pencil being good.

But if I know that the pencil is not going bad immediately, the chances the pencil is good is nothing but 10 by 20, why, because if it is not immediately going back, I have, I know that I might have chosen one of these are one of these pieces, and I have total of 20 pieces. So, my chance that the pencil is good is actually 10 by 20. So, this is an example of conditioning probability.

So, this cross B is essentially the event that the pencil does not go back immediately, and A is the pencil being chosen is good or bad. So, let us say in this case A is the pencil being chosen is good, pencil is good. So, probability of A is 10 by 25, but probability of A given B is probability of A B by probability of B so then I have a different choice. Let me come to another example.

(Refer Slide Time: 05:41)



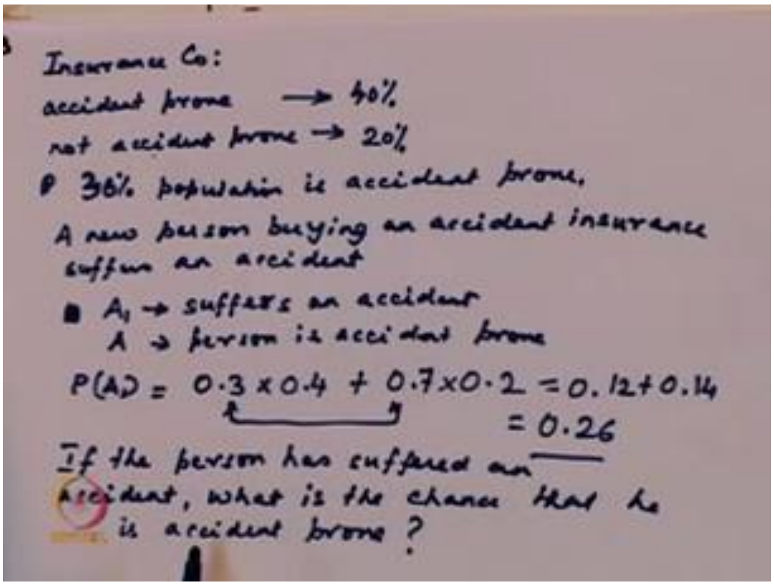
Imagine that A person X works in a bank, and he knows about this information that the. So, he walks to the bank which is far away from his place. So, this bank is located far from x's residence. So, he is eager to join a bank which is close to his home, and he has gotten this piece of information that the bank is considering opening a, you know a branch. So, bank you wants to open a branch, which is close to x's home.

So, this, and this is let say chance is forty percent. The branch is opened close to xs home is forty percent, and if the bank does open a branch close to x's home, let say the chance that X becomes the branch manager, this is 60 percent. So X becomes the branch, manager 60 percent if this particular branch is opened. So, I want to ask the question, what is the probability? So, you want to compute the probability, that X is branch manager at the newly opened branch, near his-her home. So, I want to calculate this. So, this would mean.

So, let say if I define the event A the branch opens, and B is X is chosen manager, branch manager. So, what we have to calculate here is probability of A and B so that X is branch manager at the newly opened branch near is home, which means that the branch has been opened and he is the manager there. So, how will I use conditional probability to solve this problem? So, I know probability of A given B is probability of A B by probability of B.

So, or I can write probability of B given A is probability of A B by probability of A. So, from this expression I can write P A A B is simply equal to P B given A into probability of a. So, this 6ty percent, this statement is nothing but probability of B given A, which is equal to 60 percent and probability of A is 40 percent. So, you get answer of 0.6 into 0.4 equal to 0.24. Let us revisit the problem that we had posed in last class.

(Refer Slide Time: 09:58)

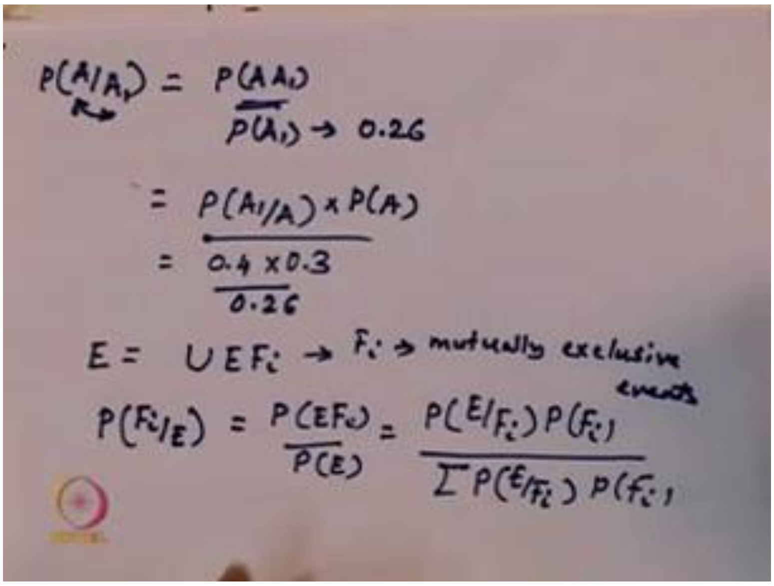


So, we had an insurance company, which labels every person as accident prone or not accident prone, these people have a chance forty percent chance of suffering an accident, and these people have 20 percent chance of suffering an accident. So, if my population, if forty percent of my population, sorry if 30 percent of my population is accident prone, I want to calculate. So, 30 percent of my population is accident prone, I want to calculate a new person buying insurance, suffers an accident. So, we have to define an event e, let say we define an event A 1, this suffers an accident, person suffers an accident, and A is the person is accident prone.

So, I want to calculate probability of A 1, which I can write as person is accident prone. So, which is 0.3, 30 percent of my population is accident prone, and with how much probability they will have an accident is 0.4, and for the remaining 70 percent. So, this and these are complimentary events, the chance of having an accident is 0.2. So, this gives me an answer of 0.12 plus 0.14 equal to 0.26. Let me modify this question and ask. So, modify my question and ask that if. So, if the accident has already happened. So, if the person has suffered an accident, what is the chance that he is accident prone? So, I want to have the reverse question, that if I know that the person has suffered an accident, what is the chance that the person is actually accident prone.

So, again in this for solving this problem also, we have to make use of conditional probability and we can write probability. So, essentially we have been asked to compute probability of A given A 1 right.

(Refer Slide Time: 13:27)



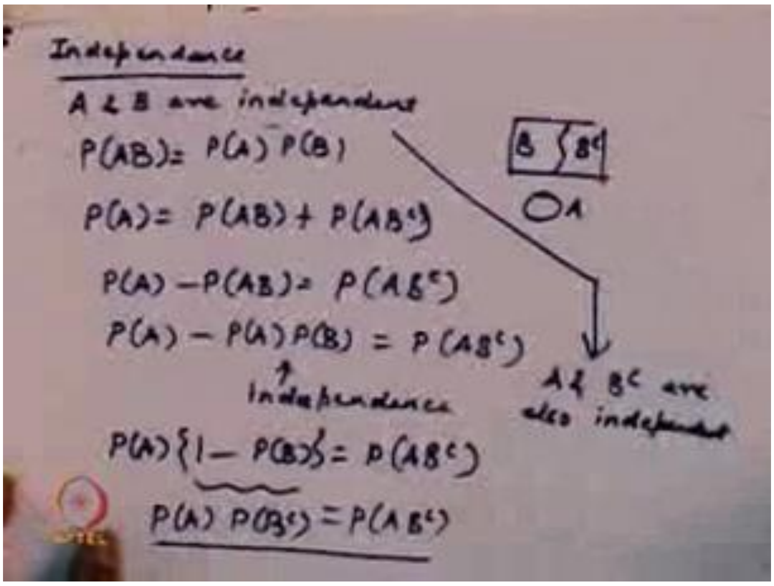
So, probability of A given A 1; so person is accident prone, given that he has an accident, I can write it as probability of A A 1 by probability of A 1, of probability of A 1 we have calculated as 0.26, what do I do with probability of A given A 1. So, I can rewrite this expression as probability of A 1 given A into probability of A. So, this is 0.4 into 0.3 by 0.26. So, this is your final answer. So, this is point. So, its 0.4 into 0.3 by 0.26 is your final answer. So, in order to calculate probability of A given A 1, you can convert the joint or intersection probability into A reverse conditional probability statement of A 1 given A and into.

So, this is how you can change the extent of conditioning of your statement you will have one more extension of base theorem; so in the generic case, when E is union of E F i. So, in order to generalize this particular finding, we can write probability of F i given E is equal to probability of E F i by probability of e, is equal to probability of E given F i into probability of F i Y summation of probability of E given F i into probability of F i.

So, here also we have generalized this finding if you have. So, here the assumption is F i are mutually exclusive.

So, you can write probability of F i given E is equal to probability of E F i and then this is nothing but each of these conditional probabilities added together, because these F i are mutually exclusive events. I come to one more concept of independence.

(Refer Slide Time: 16:10)

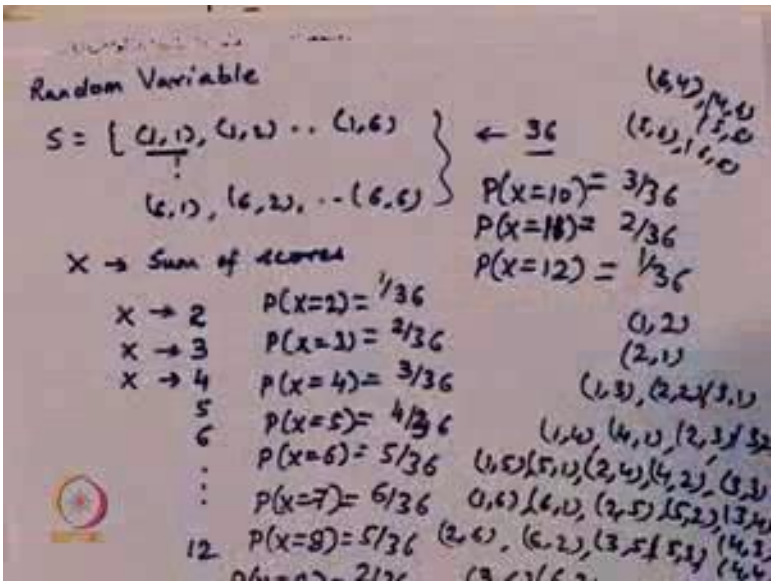


So, when 2 in the events A and B are independent. if A and B are independent, we call them as independent they can I can write probability of A B is equal to probability of A into probability of B So, because there is no dependence of A on B on vice versa, you can separate it out into 2 individual probabilities itself. Now what it follows if A and B are independent. Now I can write probability of A as probability of A B plus probability of A B complement right. So, this is your event B this is your event B complement, and this is your event a. So, they are independent. So, I have drawn them separately, but I can clearly write that P of A, p of A is P of A B plus P of A B component.

Now, I can take it in the other side, I can write P of A minus P of A B is equal to probability of A B complement. This I can simplify if A and B are independent equal to P of A B complement. So, I have invoked independence here. So, I can write this component P A into 1 minus P B equal to P A B compliment, and what is 1 minus P B its simply P A into P B compliment, equal to probability of A B complement. So, what this tells you, which means that if A and B are independent. So, we have this expression which is probability of A into B complement is equal to probability of A into probability of B complement. So, which means that if A and B are independent, A and B complement are also independent.

So, this implies A and B complement are also independent. So, this completes our initial description of probability, and conditional probability. I would next start with a concept of random variable. So, imagine I am tossing a coin twice right, or am I am rolling 2 2 coins 2 dice together.

(Refer Slide Time: 19:07)



So, in that case my sample space is 1 1 1 2 so on and so forth, up to 1 6 6 6, and as I have discussed before we have 36 different possibilities, but instead of be interested in just this combination alone, let say I can decide a variable x, which is the sum of squares sum of scores if that was possible, then what are the values of X which are possible X can be equal to minimum value is 2, I can have 3, I can have 4 5 6, maximum value of 12. So, X is a random variable this is an example of a random variable, and we have defined random variable in this particular case of sum of scores.

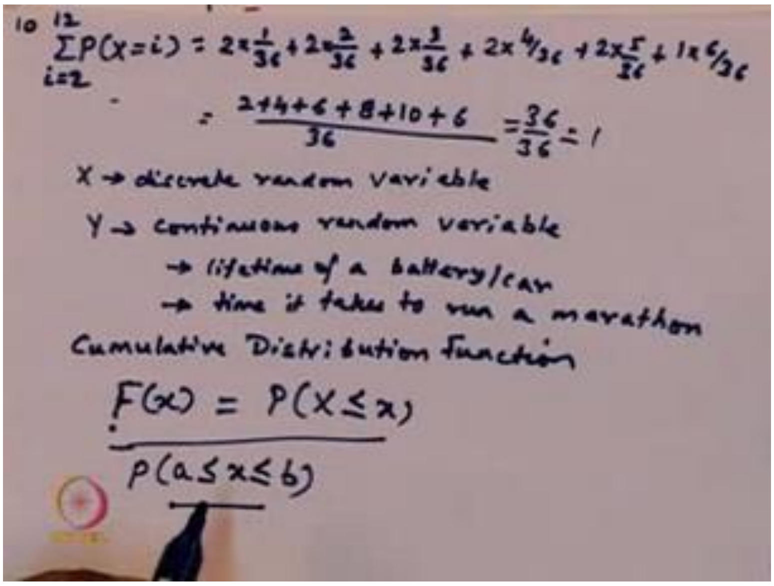
So, I can actually compute the probability of each of these events. So, I can compute probability of X equal to 2 is nothing but 1 by 36, because you have only this particular combination which gives you a value of 2. What is the probability of X equal to 3? So, what are my combinations either I have 1 2 or 2 1 probability of X equal to 3 is 2 by 36;

probability of X equal to 4 is equal to. So, I have 1 3 2 2 3 1, 3 by 36 probability of X equal to 5. So, I have 1 4 4 1 2 3 3 2. So, this is also 4 by 36.

Probability of X equal to 6 5 by this is 36, 5 by 36 probability of X equal to 7. So, I have 1 6 6 1 2 5 5 2 3 4 4 3, I have 6 by 36. probably of X equal to 8 I have 2 6 6 2 3 5 5 3 4 4 equal to 5 by 36 probability of X equal to 9, how will I get 9, I will get 3 6 and 6 3 equal to 2 by 36 probability of X equal to 10 is. So, I have 6 4 4 6 5 5 3 by 36, X equal to 11. 11 is equal to, is simply equal to. So, I have 5 6 and 6 11 6 5 2 by 36 probability of X equal to 12th is equal to simply 1 by 36. So, you have the following possibilities you have, lowest probability is 1 by 36 for X equal to 2 and X equal to 12, highest is for X equal to 7 which is 6 by 36, and then you have 5 is equal to 9 3 6, actually this is also 4 by 36 this is a symmetric distribution what you see your P is that X equal to 7 6 by 36, then you have 5 by 36 of 4 by 36 and 3 by 36 2 by 36 and 1 by 36.

So, if I add these probabilities up, what will I get? So, I have summation probability of X equal to I I equal to 1.

(Refer Slide Time: 24:11)



So, you have values between 2 and 12th is nothing but 2 into 1 by 36 plus 2 into 2 by 36 plus 2 into 3 by 36 plus 2 into 4 by 36 plus 2 into 5 by 36 plus 1 into 6 by 36. If I make the 36 common I have 2 plus 4 plus 6 plus 8 plus 10 plus 6. I have 6 plus 6 12 and 8 is 20 30 36. So, what you see is, for random variable. So, X in our case is the discrete random variable, and I can determine the summation of probability of X is equal to I simply, but is equal to the entire sample space of probability equal to 1. So, you can define any random variable, and in this particular X is a discrete random variable. I can also think of a continuous random variable, which is nothing but for example, I can have lifetime of a battery or car, or time it takes to run a marathon and so on and so forth.

So, based on the probability I can define what is called a cumulative distribution function, and it is typically reference written as F of x, is defined as priority of X less or equal to x. So, using this particular definition I can find up. So, this applies for both discrete and continuous random variables. So, I think I will stop here for today, and in the next class will see how we can determine. So, from F of X, I can find out the probability of let say X lying in a given range and so on and so forth.

With that I thank you for your attention and will continue in next class with little bit of probability and will get started on random variables and its application.

Thank you.