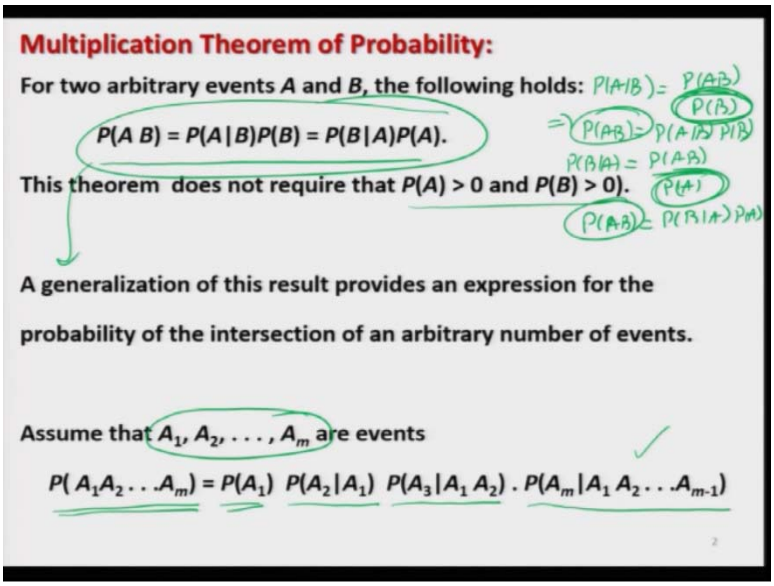
Essentials of Data Science with R Software-1
Professor Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur
Lecture 18
Multiplication Theorem of Probability

Hello friends. Welcome to the course Essentials of Data Science with R Software 1, in which we are trying to understand the basic concepts of probability theory and statistical inference. And we are continuing with probability theory concepts. So, now you have learned simple probabilities, conditional probabilities and you also have seen that they have a close relationship with the definition of relative frequency.

And that definition is going to help us when we want to compute it through the software. And you can see that simple probability and conditional probability they are going to be used under different types of condition. But the bottom line is, you are the one who is going to take a call whether you want to compute the simple probability or a conditional probability.

So, now continuing on the same lines, we have one result which is called as multiplication theorem or probabilities. And this helps us that when we are trying to deal with more than one condition more than one events, and we want to compute a particular type of probability. So, in this lecture, I am going to talk about this multiplication theorem of probability and I will try to take some examples through which I can explain you that how the things are going to work. So, let us begin our lecture.

(Refer Slide Time: 01:35)



So, now you know that in the last lecture, you have done the conditional probability, P(A|B), which is equal to P(A B) that is A ∩ B divided by P(B). So, this implies that P(A B) is equal to P(A|B) into P(B). And in case if you try to interchange this P(A |B) by P(B |A), so this will become here P(A B)/P(A), because P(A ∩ B)and P(B ∩ A), they are the same thing.

So, from here you can see here that P(A B) once again comes out to be P(B|A) into P(A). So, now, if you try to see both these values, they are the same. So that that is what I have written here, the P(A B) is equal to P(A|B) into P(B) is equal to P(B|A) into P(A).

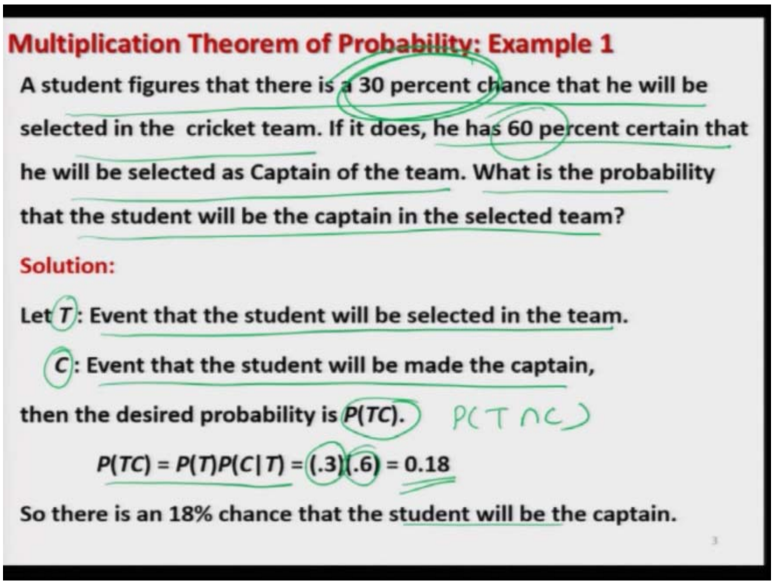
And one thing you have to keep in mind that when we try to define the probability, or the conditional probability here, you have to assume that this P(B|A), they are greater than 0 but when you are trying to define this statement, then you do not need the condition that whether P(A) is greater than 0 and P(B) is greater than 0, but obviously, means, in practice you will be always dealing with a case where the probabilities are always going to be greater than 0.

So, now in case if you try to make this result more general, then how to get it done. This result is trying to express the probability of intersection of two event that when the probability that A and B occurs together in terms of the conditional probability. So, now in case if I try to extend it that instead of here only two events A and B. Suppose, we have here a small m number of events A_1, A_2, A_m. Then the P(A_1 ∩ A_2 …∩ A_m). how this can be written in terms of this conditional probabilities given different types of condition.

So, this can be written as P(A_1) into P(A_2) conditioned on A_1 P(A_3) conditioned on two events A_1 and A_2 and up to this will continue here up to here P(A_m | A_1 A_2 … A_{m-1}). So, now, this is actually called as the multiplicative theorem of probability.

And this is very useful when you are trying to do different types of computation, although we are not considering here different types of Bayesian computations, but this Bayesian analysis and related Bayesian computation they are very important nowadays. And this theorem is going to help in those cases. That is why I want to make you understand here.

(Refer Slide Time: 05:02)



So, now let me try to take here some example to explain you the application of this multiplication theorem of probability. So, the first example which I take here is that a student figures that there is a 30 percent chance that he will be selected in the cricket team. And if it does, he has 60 percent chances that he will be selected as the captain of the team.

Now, the question is, what is the probability that the student will be the captain in the selected team? That is a very obvious question these types of incidents will happen in real life very often, and in data sciences also these types of things will happen where you would like to compute such probabilities. And my idea of taking this example is to make a one-to-one correspondence with this example with something which you will be doing as a data scientist in real life.

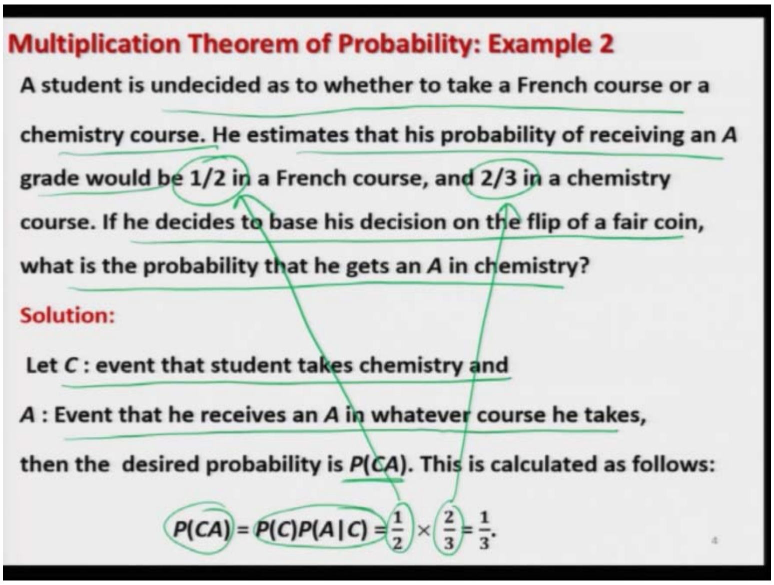
So, in order to solve this problem, let us define here two events, one is here capital T, which is indicating the team and another here is C, which is indicating the captain. So, let a capital T be the event there the student will be selected in the team and capital C be the event that the student will be made the Capitan.

Now, what you want? You want that the student is selected in the team and the student is made the captain. So, in this case, we are interested in the P(T ∩ C). So, now how to obtain it? P(T ∩ C) can be written as P(T) into P(C) condition on the that the student is selected in the team.

So, this will be P(C given T. And now, both these probabilities are known to a P(getting selected into the team is just 30 percent, which is given here. So, it is 0.3, and the and the probability that that the student was selected in the team and become the captain this is 60 percent. So, this is these probabilities, C given T is 0.6. And this comes out to be 0.18. So, there is an 18 percent chance that the student will be selected as the captain after getting into the team.

Now, even its quite occur in real life that you want to know something after something has happened. So, this type of conditional probability will help you through the multiplication theorem to give you the desired outcome.

(Refer Slide Time: 07:31)



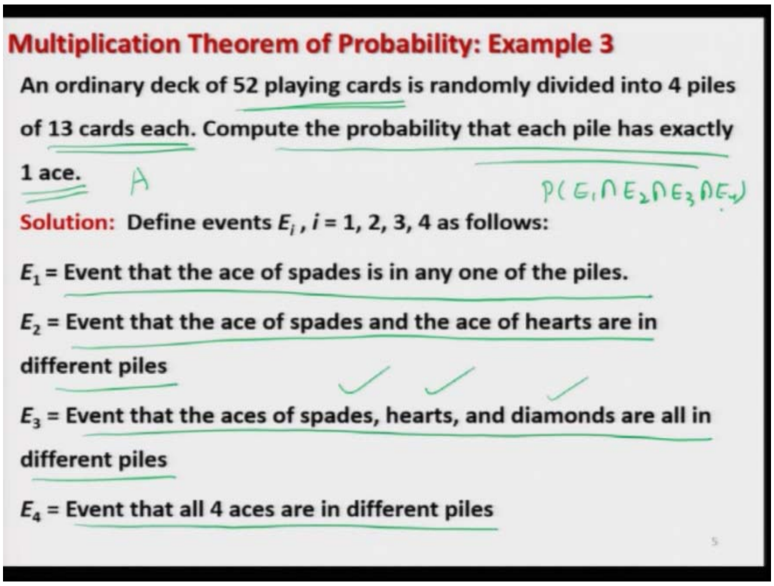
So, now, let me take care of one more example. Suppose there is a student, who is undecided to take whether a French course or a Chemistry course. Well, this happens, and in these cases, the student will always try to see that which of the course is going to give him or her a better grade. So, similarly the student also estimates his probability of receiving an A grade in Chemistry and French course, and he finds that that find that getting an, getting a grade A in the French course is 1/2 and in the Chemistry course, it is 2/3.

So, now if the student decides to base his decision on the flip of a fair coin, that he says that, I will toss a coin, if it is comes head, then I will take this subject, if it come tail then I will take this subjects. So, in case if the students decided to base his decision on the flip of a fair coin, then we want to know what is the probability that he gets an A in Chemistry.

So, now you can see here that this event is going to be based on the outcome of another event of flipping a coin, whether the student gets a head or a tail. So, now let us try to define here two events. Let C be the event there the student takes Chemistry and A is the event that he receives an a grade A, in whatever course he takes either French or Chemistry. Then in this case, the desired probability, what we want to know is the P(C ∩ A) that the person takes Chemistry as the, as well as he gets a grade A.

So, this is calculated by using the multiplication theorem of probability as P(C ∩ A is P(C into P(A |C), which is and the P(C) is 1/2. This is given here and the P(A | C) is here like this. So now this will come out to be a 1/3. So now based on this, there are nearly 33 percent chances that the student will get grade A in the Chemistry.

(Refer Slide Time: 09:45)

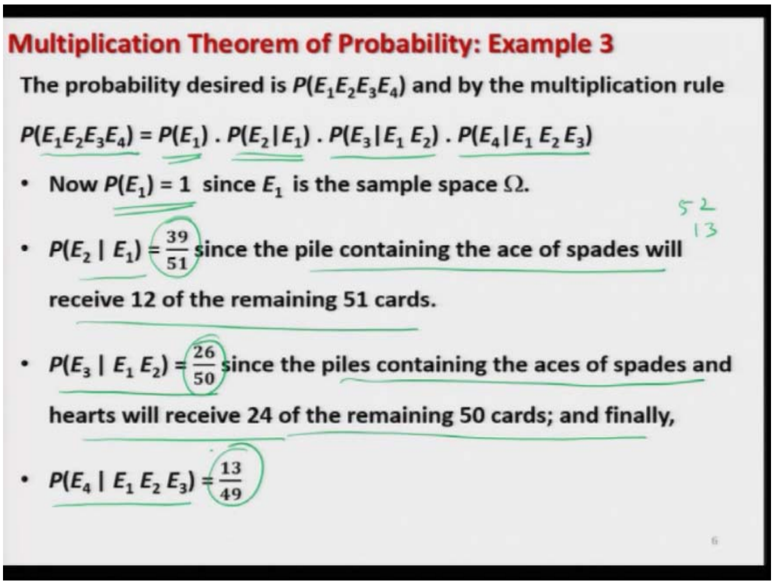


Now, let me try to take here one more example. And this example is related to the playing cards. So, we know that in the box of playing cards, there are 52 cards. And suppose we tried to make 4 piles of these cards and they are equally numbered. So, every pile will have 13 cards each. Now, we want to compute the probability that each pile has exactly one Ace. Ace, you know that in the cards there is a card with a symbol A.

Now, we define an events say E_i, E_1, E_2, E_3, E_4 depending on different possibilities as follows. So, let E_1 be the event that the ace of spades is in any one of the piles, E2 that the ace of spades and the ace of hearts are in different piles, E_3 is the event that the aces of spades, hearts, diamonds are all in different piles, E_4 is the event that four aces are in different piles. You know that in the playing cards, we have different types of symbols spades, hearts, diamonds etc. So, that is related.

So, now based on this, we want to compute the probability that each pile has exactly one ace. So that means we are interested in E_1, $P(E_1 \cap E_2 \cap E_3 \cap E_4)$. So, now, so, this probability can be obtained using the multiplication theorem of probability as follows.

(Refer Slide Time: 11:30)

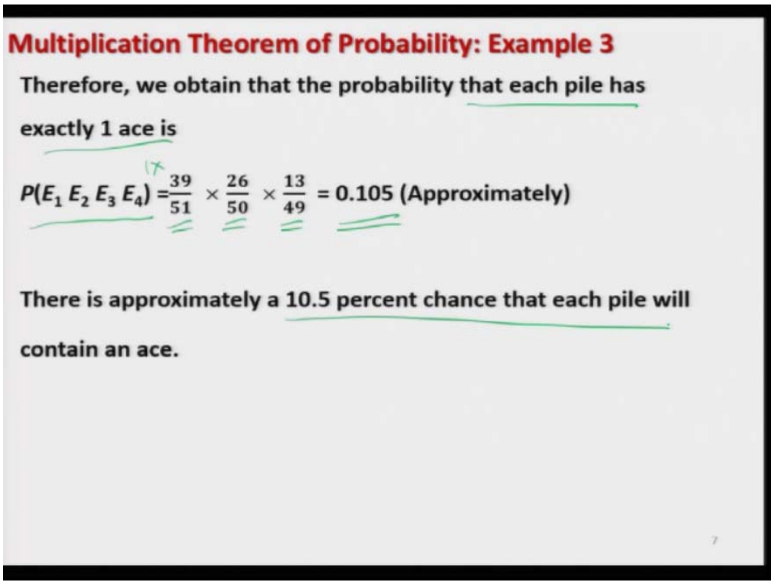


Now, I can write down here $P(E_1 \cap E_2 \cap E_3 \cap E_4)$ as $P(E_1)$ into $P(E_2)$ which is conditioned on E_1, $P(E_2)$ which is conditioned on two events E_1 and E_2 and $P(E_4)$ which is conditioned on three events E_1, E_2 and E_3.

Now, I try to compute these probabilities based on the different number of cards in the this group of cards. So, now, the $P(E_1)$ will be equal to 1, why? Because if you try to see you are trying to define here E_1 as the event that ace of spade is in any one of the piles. So, definitely you are trying to divide all the cards into 4 piles towards, so definitely it is going to be there in one of the pile. So, in this case, the probability becomes here 1.

Now, the $P(E_2 | E_1)$, this will be 39/51, why? Because the pile containing the ace of spades will receive 12 of the remaining 51 cards, because you have total 52 cards and you want to and you have a pile of 13 cards. And similarly, if you want to find out the $P(E_3 | E_1, E_2)$ this will come out to be 26/50 because the piles containing the aces of spades and heart will receive 24 out of the remaining 50 cards. And finally, this $P(E_4 | E_1, E_2, E_3)$ this will be coming out to be 13 upon 49.

(Refer Slide Time: 13:12)



So, now, in case if you try to just substitute all these values over here, then we can find out the probability that each pile has exactly 1 ace $P(E_1 \cap E_2 \cap E_3 \cap E_4)$ as 39/51 into 26/50 into 13/49 and obviously, 1 into these things. So, this is 0.105 approximately. So, there is approximately a 10.5 percent chance that each pile will contain an ace.

So, now, we come to an end to this short lecture. Well, this lecture was short and I have taken a very simple example, to explain you and I have taken a very simple result to explain you various types of property that can be computed from it, but this is a very important result I am telling you.

So, now as always I will request you that you try to take some example from your assignment, from your books and try to practice them. Once you have this practice that how to take a call that where you have to found, this type of probability and how are you going to find out these probabilities that will help you a lot in the Decision Sciences or Data Sciences.

And in Data Sciences, we have one more very popular approach this is called Bayesian approach. And this Bayesian approach is mostly now based on computing, computers. And because of computing only this, Bayesian analysis has become very popular, is after the advancements in the computing and computers. So, this theorem is going to help you a lot in case if you are trying to learn the Bayesian inference. So, with this aim you try to practice it, try to learn more and I will see you in the next lecture. Till then, goodbye.