IIT Madras

ONLINE DEGREE

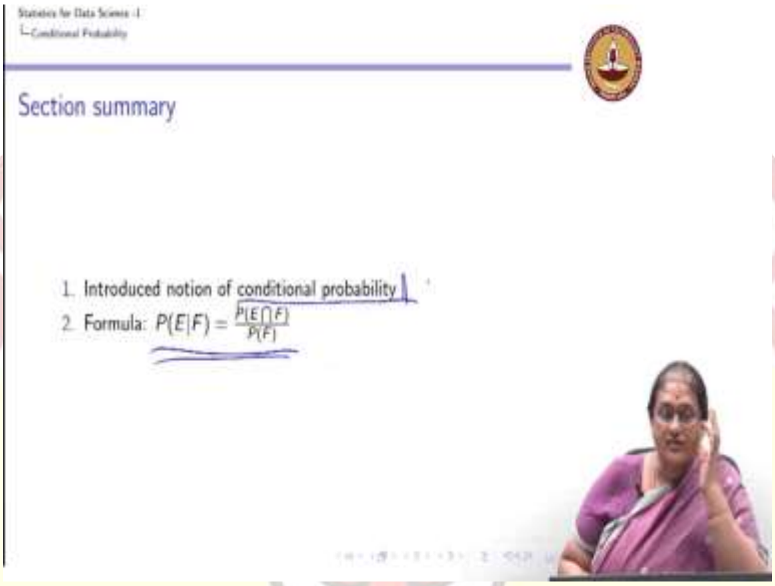
Statistics for Data Science – 1

Professor Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras

Lecture No. 7.3

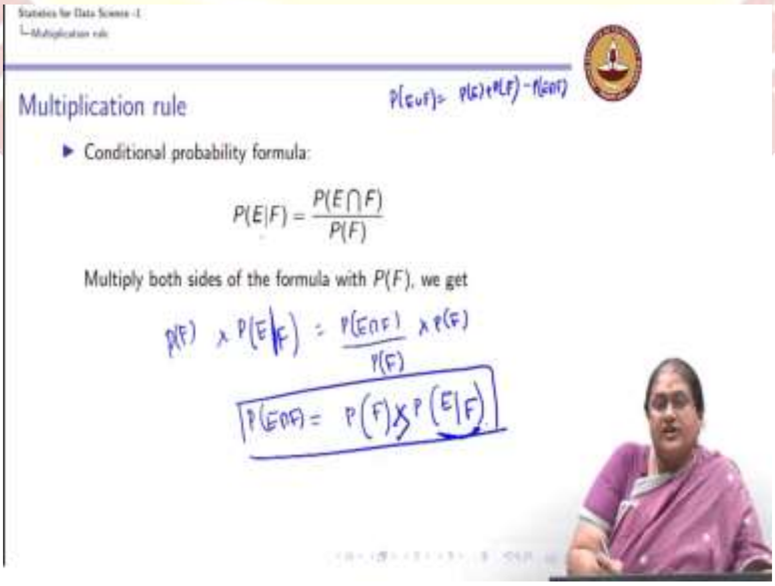
Conditional Probability - Multiplication rule

(Refer Slide Time: 0:14)



So, what we have introduced is the important notion of conditional probability which takes care or discusses about how to compute probability of event's condition or given the partial information that some other event has occurred and this is the formula for conditional probability, $P(E|F) = \frac{P(E\cap F)}{P(F)}$ with $P(F) > 0$.

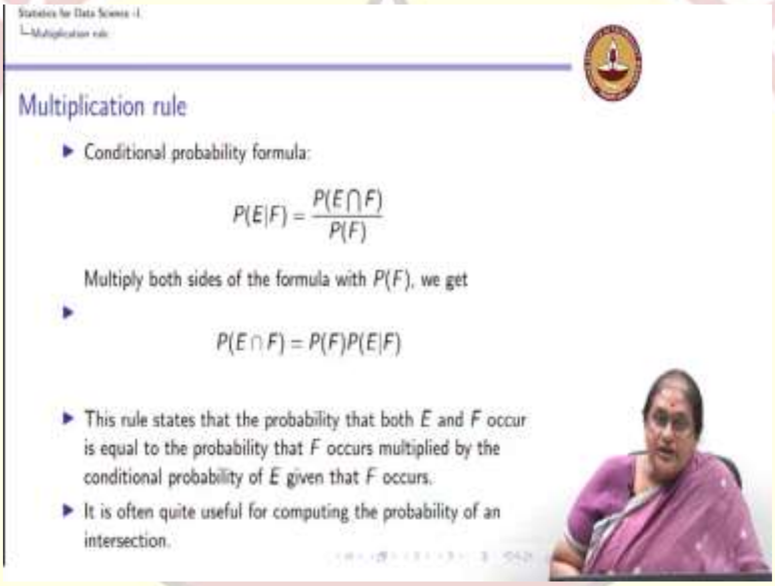
(Refer Slide Time: 1:00)



Now, from this formula which we have derived or we have given for conditional probability, we derive an important rule which we refer to as the multiplication rule. Recall that the probability, conditional probability formula tells, $P(E|F)$, again I repeat that this is not E divide by F, it is E given F or E conditioned on F is $\frac{P(E \cap F)}{P(F)}$.

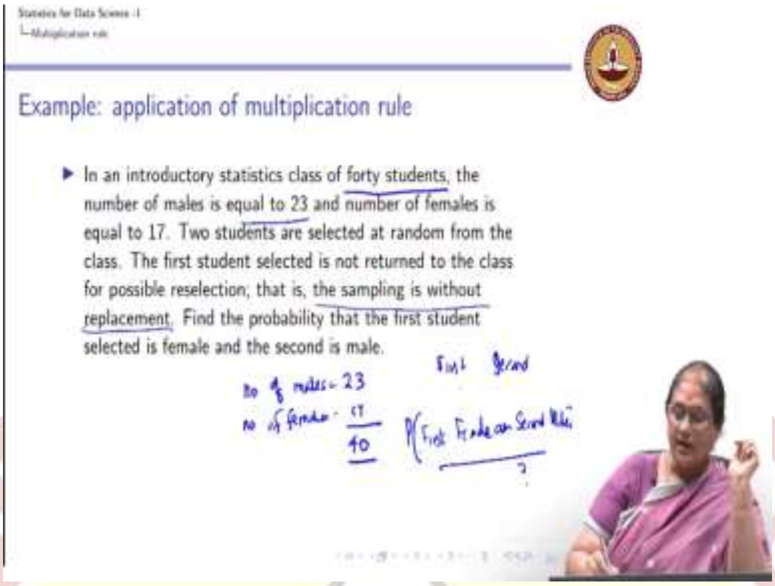
I can multiply both the sides with $P(F)$, so I get $P(E \cap F)$. So, if I multiply both the sides with $P(F)$ because $P(F) > 0$, I will get $P(E \cap F) = P(F)\, P(E|F)$. So, recall the additional rule gave us $P(A \cup B)$ or $P(E \cup F) = P(E) + P(F) - P(E \cap F)$. This is what our addition rule stated. Now, the multiplication rule states that $P(E \cap F) = P(F)\, P(E|F)$.

(Refer Slide Time: 2:26)



So, this multiplication rule actually states that the probability of both E and F occurring together, that is the probability of intersection is equal to that the probability that F occurs that is this multiplied by the conditional probability of E given F occurs which is this. So, this multiplication rule is useful to compute the probability of intersection of events.

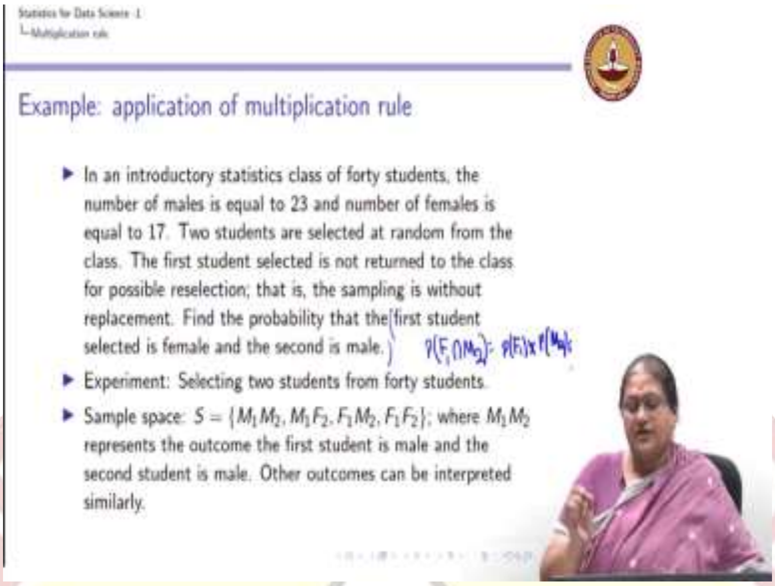
(Refer Slide Time: 2:58)



So, now let us look at an example of application of the multiplication rule. So, in an introductory statistics class of 40 students, so total number of students in the class is 40, the number of males = 23 and number of females = 17. And total number is 40 students. I am selecting 2 students at random and the way I am selecting these 2 students is I have a first student and a second student, if I am selecting a first student that student is not available for selection. In other words it is called a sampling without replacement.

So, if I select the first student, this student is not available for selection for the second student, so this is called a sampling without replacement, I am not having that student available. So, the first student and the second student. So, what are we asking? We are asking for probability that the first student who is selected was a female and second student is a male. This is the question we are asking.

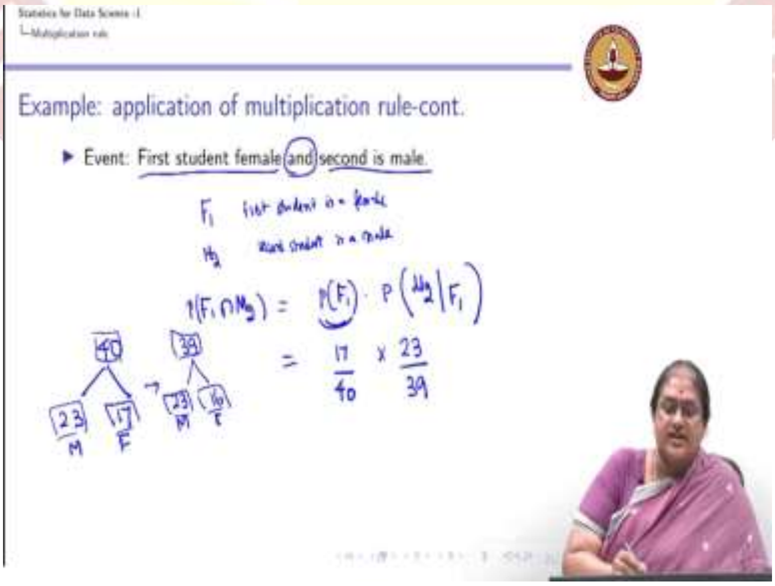
(Refer Slide Time: 4:35)

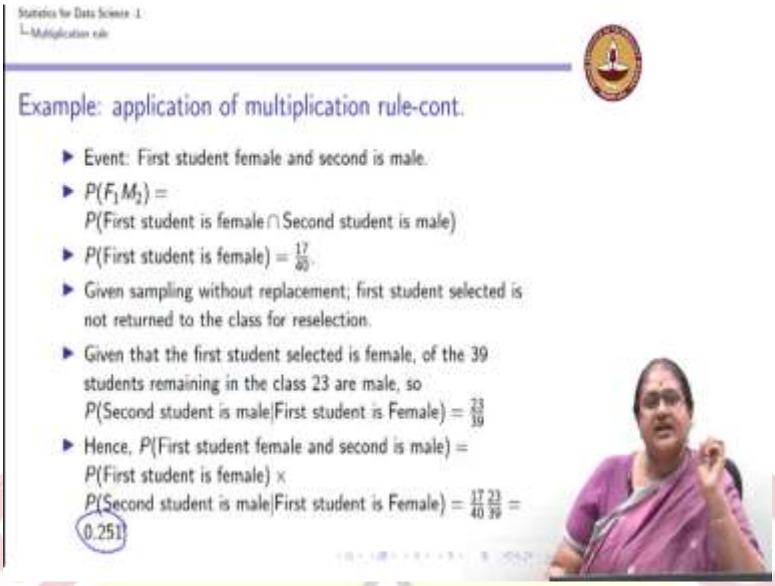


So, now if I put this in the framework which we know the random experiment here is to select 2 students and how are we selecting these 2 students, we are picking up one student, keeping him aside, then we have the remaining 39 picking up another student. So, the random experiment is selecting 2 students from the 40 students. I can list by sample space in the following way M_1M_2, M_1F_2, F_1M_2 and F_1F_2, where M_1M_2 refers to the first student being a male and the second student also being the male.

F_1M_2 is first student is a female and the second student is a male. M_1F_2 is a first student is a male and the second student is a female and F_1F_2 is the chance that both the first and the second student selected are female.

(Refer Slide Time: 5:37)





Now, from this event, I am defining the event that the first student is a female and second student is a male. So, F_1 represents the outcome that the first student is a female, M_2 is the outcome second student is a male. Now I want to know $P(F_1 \cap M_{2)}$ because here it is asking what is the chance that the first student is a female, and second student is a male.

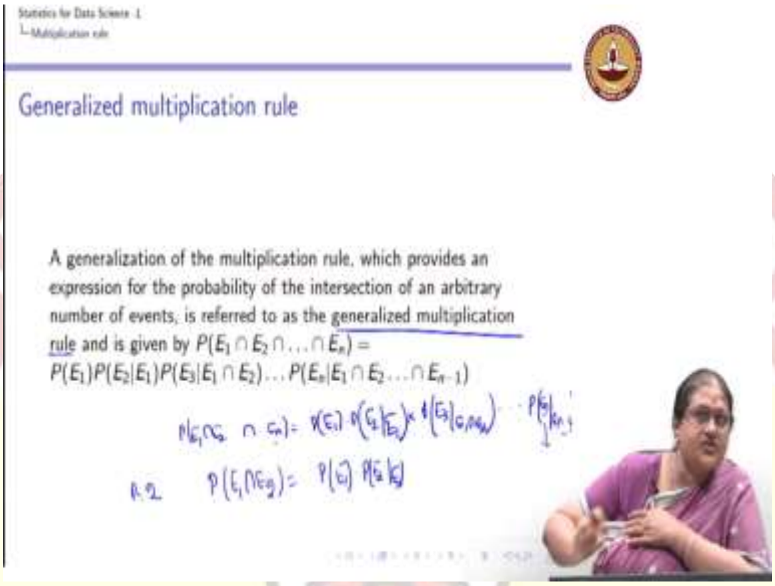
Now, if I apply the multiplication rule, know that this is probability of first student being a female into probability of the second student being chosen is a male conditioned on the fact that the first student was a female. Now if you look at what is the probability of the first student being a female, I have 40 students of which I have 23 of them who are male and 17 of them who are female.

So, the chance of my first student being a female is 17/40, because I have 17 females and 40 total student, so the chance of me selecting a female out of 40 students is 17/40. Now once I have selected this female student, the total number of students become 39 and this I have already selected a female so I have only 16 females remaining, but I have 23 males remaining.

So, given that there are 20, the first student was a female, now the chance of me finding a male reduces to 23/39 because the total number of students is now 39 with 23, this distribution or this separation or this classification that 23 are male and 16 are female. So, the chance of the first student being female and second student being male is $\frac{17}{40} \times \frac{23}{39} \approx 0.251$.

So, this is what we referred to as the multiplication rule and we have seen an example where the multiplication rule has been used to compute the probability of a intersection. So, what we need to understand is translate the problem in terms of the events so that we understand what is the event that we are seeking. For example, here the event that we are seeking is female 1 and the second is a male, it is expressed as an intersection of the events and once we have expressed it as an intersection of events, we applied the multiplication rule to get hold of the probability of the event.

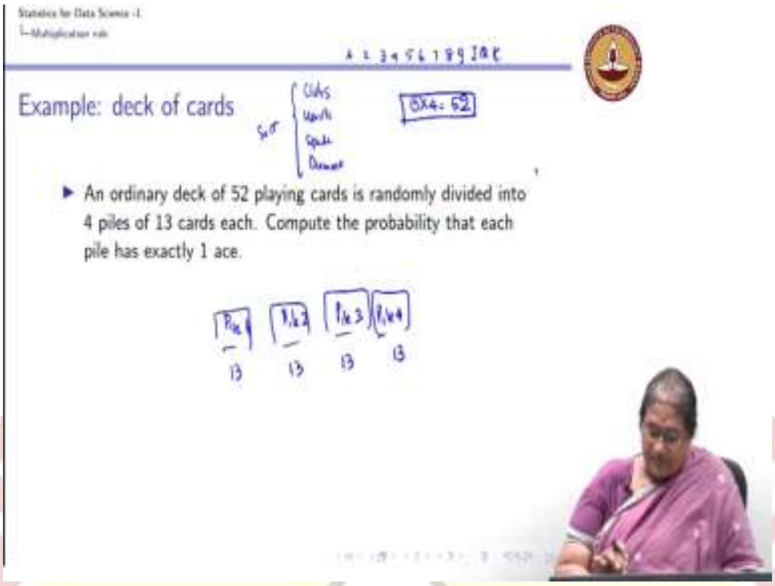
(Refer Slide Time: 9:06)



So, this multiplication rule can be generalised to arbitrary number of events. Suppose I have n events, then I can refer to that as a, and I refer to this as a generalized multiplication rule. What is the generalized multiplication rule? If I have n events, $P(E_1 \cap E_2 \cap \ldots \ldots \cap E_n) = P(E_1) \times P(E_2|E_1) \times P(E_3|E_1 \cap E_2) \ldots \ldots . . \times P(E_n|E_1 \cap E_2 \cap \ldots \ldots \cap E_{n-1})$.

So, when my $n = 2$, it reduces to what we already have is $P(E_1 \cap E_2) = P(E_1)\, P(E_2|E_1)$, because if $n = 2$, it reduces to this term just because $P(E_2|E_1)$ because E_{n-1} is E_1 which is what I already have seen. So, which is, $E_2|E_1$ which is something which we have already computed.

(Refer Slide Time: 10:26)

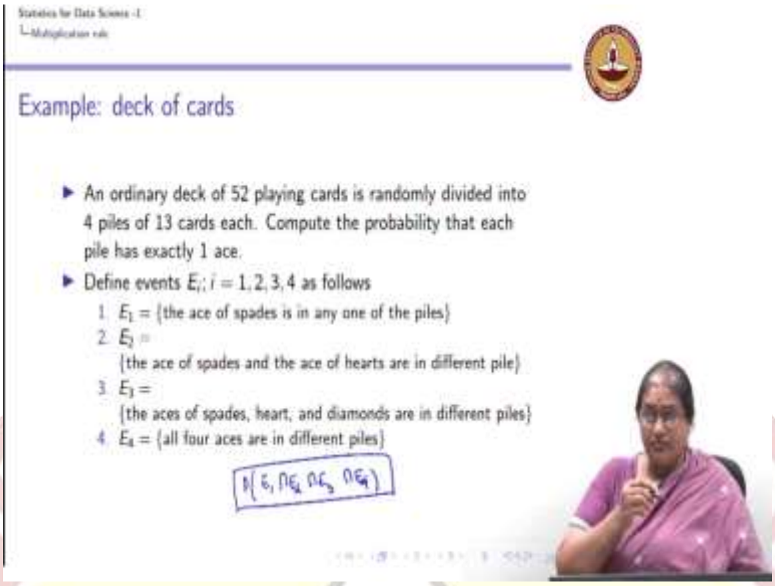


So, let us look at an example of the application of the generalized multiplication rule. Again consider a deck of 52 playing cards. Now all recall when I talk about a playing cards, I have clubs, I have hearts, I have spade and I have, so I have 13 faces with these are what I referred to as a suit of a card. So, 13 × 4, I have 52 cards. We have already seen a card example when we discussed about probability of events.

So, what we are doing from these 52 cards is I have picked up, I have randomly dividing this into 4 piles of 13 cards each. So, we are randomly choosing cards we are preparing pile 1, pile 2, pile 3, so I have 4 piles. So, what are we doing? We are picking up cards at random and putting them into 4 piles or 4 buckets whatever way we would like to choose, pile 3 and pile 4. So, I pick up 1 card, I put. So, I start putting, I shuffle these cards and I start putting them into these 4 buckets or 4 piles randomly. So each pile now has 13 cards.

So, now we want to compute the probability that each pile as exactly 1 ace. So, I could have a ace of hearts here, I could have a ace of diamond here, I could have a ace of spade here and a ace of club here, but I cannot have a situation where pile 1 has an ace of heart and a ace of club, ace of spade here, ace of diamond here. This cannot, this is not permitted. But I want each of these piles to have exactly one ace. How do I apply the general multiplication rule to solve this problem?

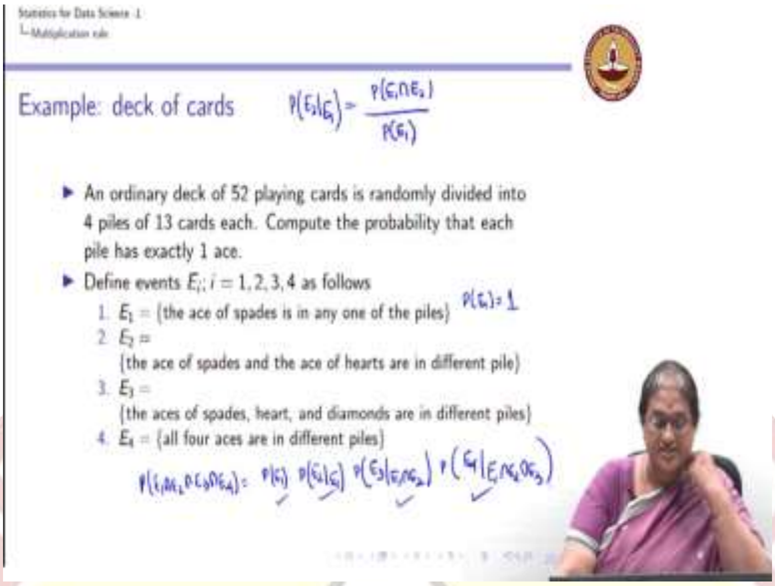
(Refer Slide Time: 12:52)



So, now let us define a few events. Now let E_1 be the event that the ace of spades is in any one of the pile. So, remember I have 4 piles. I want to define an event that ace of spades is in any one of these piles. The second event is ace of spades and ace of hearts are in different piles. There is no sanctity about starting with ace of spades, I could have started with ace of hearts and defined ace E_2 to be the event ace of hearts and ace of spades are in different piles.

The third event is ace of spades, heart and diamonds are in different piles and the fourth event is all 4 aces are in different piles. So, the question that compute that the probability that each pile has exactly 1 ace is equivalent to finding out the probability that $P(E_1 \cap E_2 \cap E_3 \cap E_4)$, this is what I want to find out that each pile has exactly 1 ace is equivalent to finding out this probability.

(Refer Slide Time: 14:22)

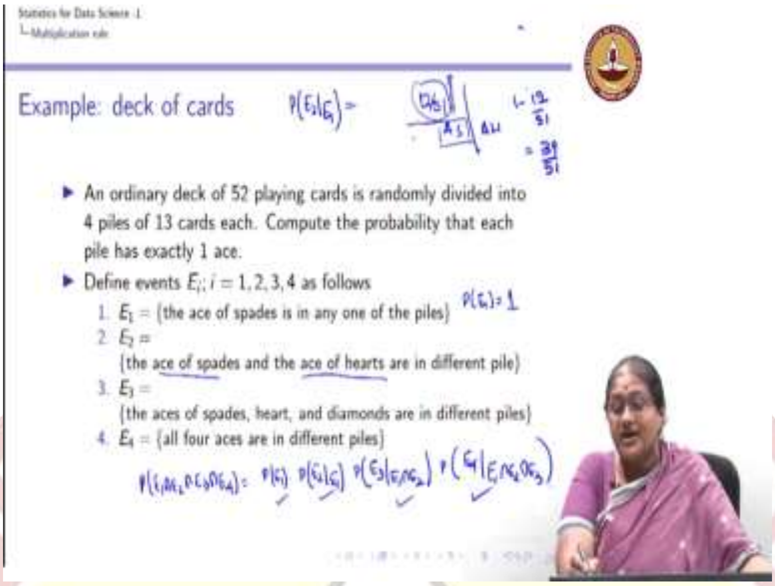


And by general rule of probability, I have that $P(E_1 \cap E_2 \cap E_3 \cap E_4) = P(E_1) \times P(E_2|E_1) \times P(E_3|E_1 \cap E_2) \times P(E_4|E_1 \cap E_2 \cap E_3)$. So, once I find out these 4 probabilities, I can answer the probability on the left hand side which is the probability we want to compute. Now let us look at each event's probability. What is event E_1?

Event E_1 is that the ace of spades is in any one of the piles. So, what does this mean that, what is the chance that I can find ace of spades in any one of the piles, I know that this is equivalent to the probability of the sample space that is ace of spade, so I have these 52 cards, I have put it in 4 piles. So, ace of spades is in this or this or this or this so the chance of finding ace of spades in any one of them is equal to 1, so I can write down $P(E_1) = 1$ or it is a sure event that is I can find the ace of spades being in any one of the piles is a sure event.

Now, let us look at $E_2|E_1$. From my definition I know $P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$. Now what is $E_1 \cap E_2$? $E_1 \cap E_2$ is the ace of spades and ace of hearts are in different piles. So, now let me look at the conditioning part.

(Refer Slide Time: 16:34)

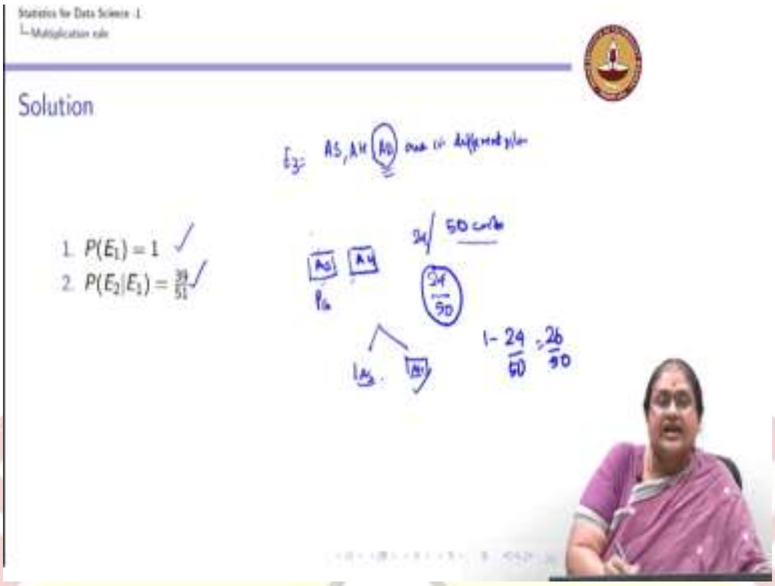


I look at E_1 and it is a ace of spade, I put that in one of the piles. Let me put it in say pile 1. Now, each pile has 13 cards, each of the piles have 13 cards, so if I put ace of spades as a one of the cards in this pile 1, the remaining 12 cards can be chosen from, so I am fixing this ace of spades, so I have 51 cards that are remaining so I can choose the remaining 12 cards in this pile of which ace of spades is there. See, I repeat, I take a pile on which I already have ace of spades, this could be any of the piles, but if ace of spades is one of the cards in a pile, I have to choose 12 remaining cards in the pile.

Now, these 12 remaining cards can be chosen from 51 cards. Now, how can they be chose? Why are they chosen from 51 cards? Because the ace of spades is not available for my choice now. And that can be chosen from 51 cards in 12/51 ways. Now I do not want ace of hearts to be one of these 12/51 choices. I do not want the ace of hearts because I do not want ace of spades and ace of hearts to be in the same pile.

The remaining 12 cards can be chosen by 12/51 ways. But I do not want the ace of hearts to be in this 12/51 ways. So the way that ace of spade and ace of hearts will be in different piles is going to be $1 - \frac{12}{51} = \frac{39}{51}$ ways. I am just using the conditional probability logic by if I am conditioning it on the fact that ace of spades is already occupying a law, the remaining 12 cards can be got in 12/51. I do not want ace of hearts to be in this choice, so the way ace of hearts cannot be in this choice is $1 - \frac{12}{51} = \frac{39}{51}$.

(Refer Slide Time: 19:12)

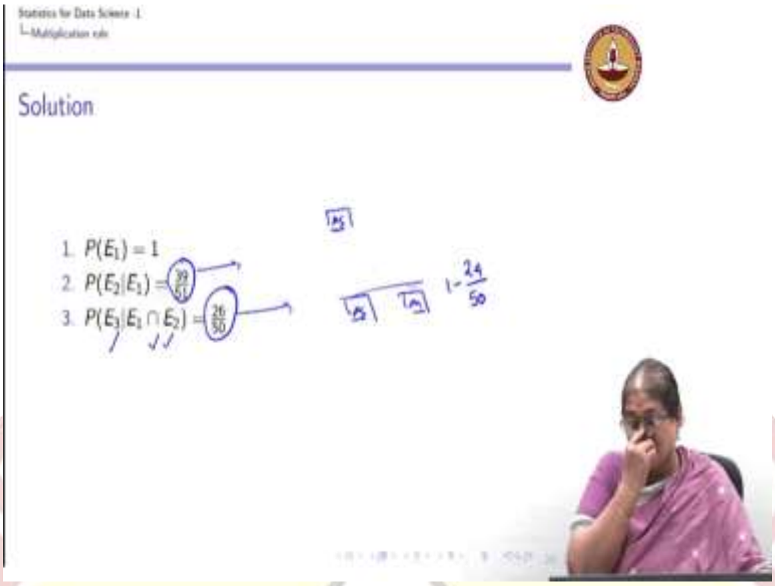


And hence, $P(E_2|E_1) = \frac{39}{51}$. Now what is E_3? E_3 is ace of spade, ace of hearts and ace of diamond are in different piles. So, I already have ace of spade in one pile, ace of hearts in one pile, so if I am fixing ace of spades in one pile and ace of hearts in one pile, I have 50 cards because these 2 cards are already fixed. Now to fill up this pile 1 say, I need a 12 cards, to fill up this I need another 12 cards. So, I need a total of 24 cards to be chosen out of 50 cards.

I repeat. I have ace of spades in one pile, ace of hearts in one pile, so if I fix ace of spades in one pile and ace of hearts in one pile, I have only 50 cards to be chosen from total 50 cards. In pile 1, I can, I need to choose 12 cards, pile 2 I need to choose 12 cards, so in total I need to have 24 out of 50 that has to be chosen. I am writing 24 out of 50 because we can interchange these choices and now 24/50 ways I have AS and A hearts are in different piles.

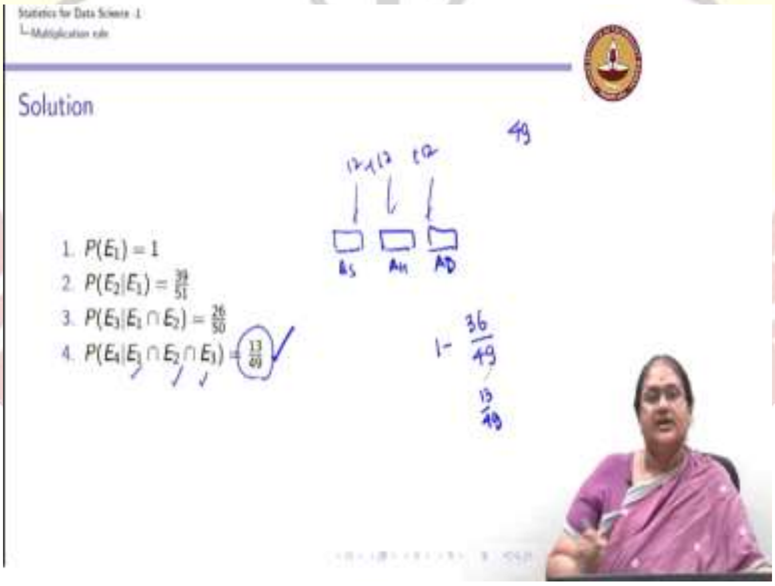
Now, I do not want A diamond to be in either of these piles. So, the number of ways ace of diamond is not in these of any of these 2 piles, I know there are 24/50 ways it can be in these 2 piles, so $1 - \frac{24}{50} = \frac{26}{50}$ ways I will have ace of diamond not in the piles in which already ace of spades and ace of hearts is.

(Refer Slide Time: 21:37)



So, $P(E_3|E_1 \cap E_2) = \frac{26}{50}$. Now similarly using the same logic so we got this 39/51 by fixing ace of spades, we got this 26/50 by fixing ace of spades and ace of hearts and we said that the remaining 24 cards that need to go into these 2 piles can be got by 24/50 ways, 24/50 ways and the way I can ensure that ace of diamond is not a outcome of this, so I get $1 - \frac{24}{50} = \frac{26}{50}$.

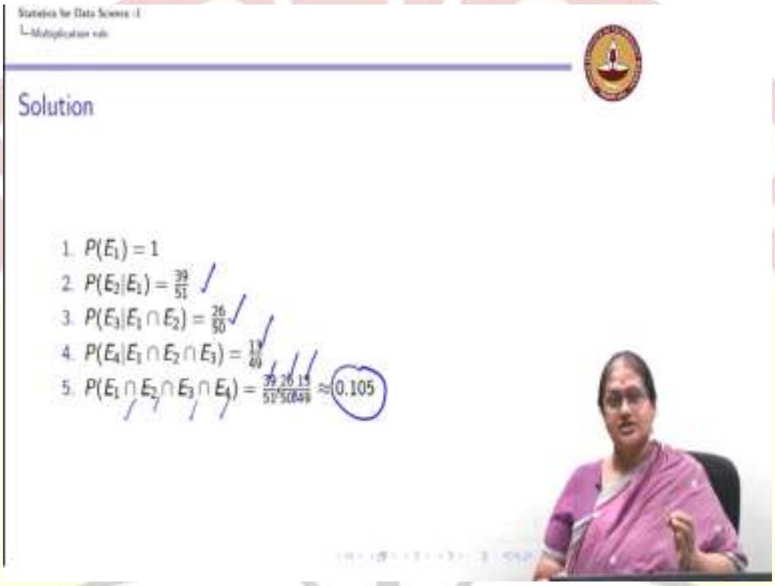
(Refer Slide Time: 22:27)



Similarly, I can argue that $P(E_4|E_1 \cap E_2 \cap E_3) = \frac{13}{49}$. Now I have 49. So, I already I am fixing ace of spades, ace of heart, ace of diamond, so I have 49 cards that are available. Again 12×3 which is 36 ways I can actually choose these cards, the remaining, so for each of these piles I have 12 + 12 + 12 which is 36 cards which can be chosen from the 49 cards in 36/49 ways.

I do not want my cards to be any one of these 36/49 choices. So, I look at $1 - \frac{36}{49}$ which is given by the number here. So, we have computed all the, whatever we needed to compute $P(E_1 \cap E_2 \cap E_3 \cap E_4)$.

(Refer Slide Time: 23:41)



We just plug in these values to get $P(E_1 \cap E_2 \cap E_3 \cap E_4) = \frac{39}{51} \times \frac{26}{50} \times \frac{13}{49} \approx 0.105$

(Refer Slide Time: 24:04)

So, what we have learned so far in this section was, what is a multiplication rule, when I have two events I know $P(E_1 \cap E_2) = P(E_1) \times P(E_2|E_1)$. We extend this to give the generalized multiplication rule which is for any arbitrary n events, it is $P(E_1 \cap E_2 \cap \dots \dots \cap E_n) = P(E_1) \times P(E_2|E_1) \times P(E_3|E_1 \cap E_2) \dots \dots .. \times P(E_n|E_1 \cap E_2 \cap \dots \dots \cap E_{n-1})$. And we further use this generalized multiplication rule to compute probability of intersection of events.