Essentials of Data Science With R Software-1.

Professor Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

Lecture 14.

Some Rules of Probability

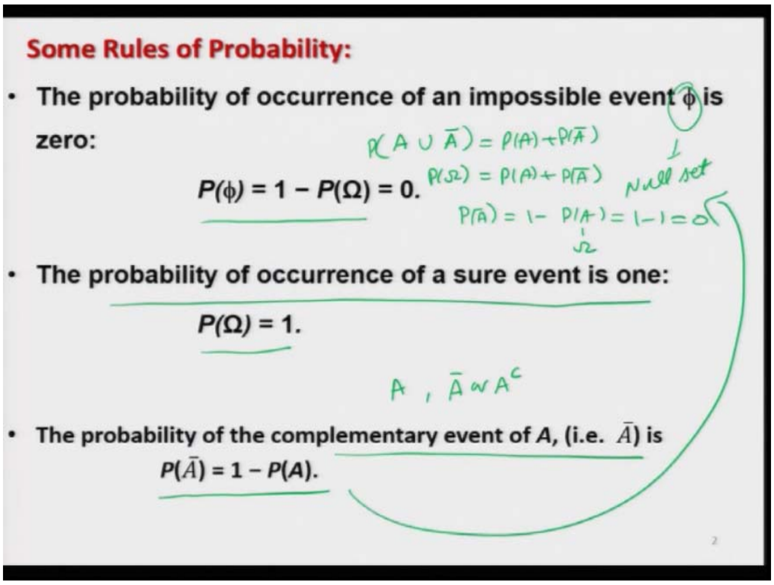
Hello friends. Welcome to the course Essentials of Data Science with R Software-1, where we are trying to understand the topics of probability theory and statistical inference. So, you can recall that in the last couple of lectures, we have tried our best to understand the basic fundamental concepts of the probability theory by introducing very elementary definitions.

And we have tried our best that whatever we are trying to study from the theory point of view, from the critical point of view, how it will actually work in real life? How this theory and numerical values are interrelated? And how the numerical values can be justified on the basis of mathematical theories?

So, now, in this course, I will try to give you some basic rules, which will help you in the computation of different types of probabilities and a very simple example to show you that how the rules what we have learned earlier and in this lecture are going to be useful for us. And my objective is this, I want to show you that, if I have a situation or a condition where I want to know something, how those things can be translated in terms of probability theory, and how they are going to interpret it.

Once you understand this thing, then I will move further. That in order to achieve this type of probability, what else you have to learn. So, we begin our this lecture and we try to understand some basic fundamental rule for the calculations of probabilities.

(Refer Slide Time: 02:00)



So, the probability of occurrence of an impossible event, how to find it out? We know that whenever we are talking of an impossible even that means, there is no event in the that set. And that set is indicated by ϕ, which is actually here a null set. So, the probability of occurrence of an impossible event that is ϕ is 0.

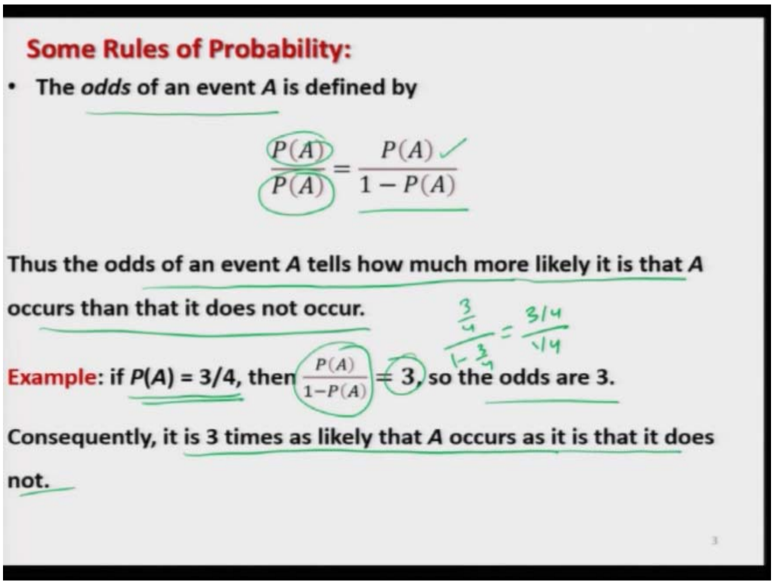
So, we always write that P(ϕ) is equal to 1- P(Ω), which is equal to 0. Because you know that A and A complement, these are two are the joint sets. So, P(A) and A complement, this is going to be P(A) + P(A^c) .

P(A ∪ A^c) is same as probability of Ω, this is a sure event, which is equal to here P(A) + P(A^c). So, we know that P(A^c) complement is equal to 1-P(A). So, now from here you can see here, when your A is Ω, that means the probability of Ω that we say that is the sure event is always equal to 1. So, 1-1 will be equal to 0. So, this is how we have got the results stated in this slide.

So, the first rule what you have to understand that the probability of occurrence of an impossible event is 0. And the probability of occurrence of a sure event is 1, that probability of Ω is equal to 1. And similarly, in case if you have two events A and A complement, or you can also denote it by A^C indicating the complement.

The probability of the complementary event of A is P(A) complement is equal to 1-P(A), that I already have shown you here. That is a pretty simple thing. So, these are very simple rules which we try to follow in general.

(Refer Slide Time: 04:07)

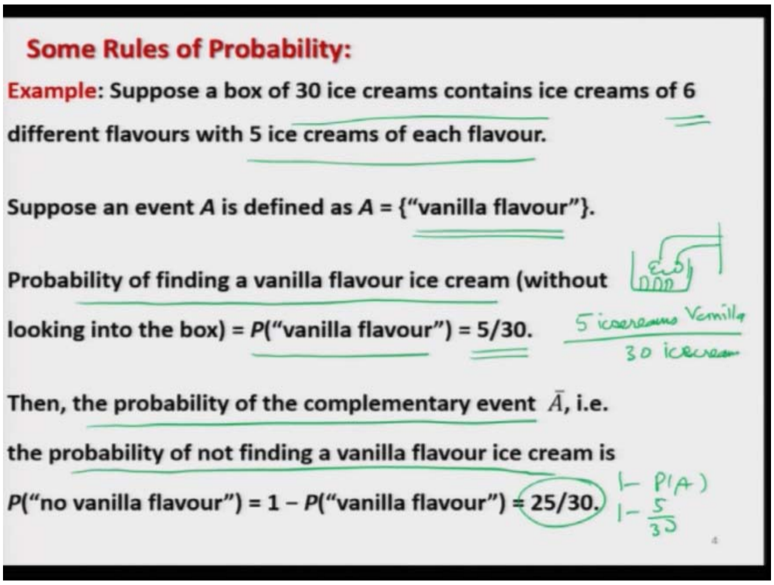


Sometimes, instead of computing directly the probability we always, we also try to actually compute the odds of an event, which is actually the function of P(A) and its complement. So, and that is defined as the ratio of P(A) and P(A^c). So, that is equal to P(A) /(1-P(A)).

The odds of an event A tells us how much more likely it is that A occurs then it does not occur, as simple as that. So, it depends on your objective, whether you wants to know the results in terms of odds or in terms of P(A). For example, if I say that if a P(A) is certified or sent that is P(A) is equal to 3/4, then the odds of A will be P(A)/1-P(A) that will come out to be here 3/4 divided by 1-3/4, which is equal to 3/4 divided by 1/4 which is equal to 3.

So, that means, the odds are 3. This means it is three times as likely that A occurs as it is that it does not occur, as simple as that, very simple interpretation. So, now, it is up to you what you really want to use in the given condition, given situation.

(Refer Slide Time: 05:37)



Now, I try to take a very simple example to explain you these concepts through it. So, suppose there is a box in which there are 30 ice creams and there are 6 flavors of the ice cream in the box and there are 5 ice creams of each of the flavor. So, altogether there will be 30 ice creams.

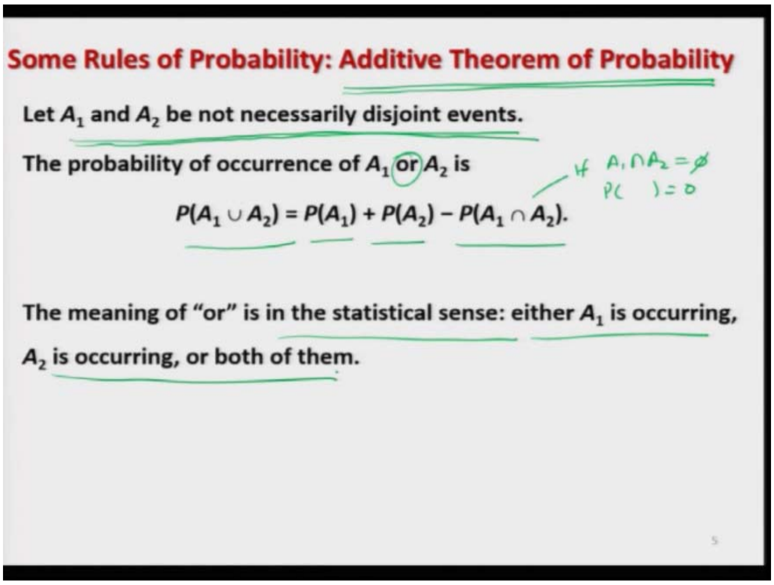
Now, we try to define an event A as vanilla flavor, well that is a flavor of ice cream. So, now, in case if I want to find out the probability of finding vanilla flavor ice cream that means, that there is a box where there are some here ice creams and you try to put your hand inside it and you are not looking into it. If you look into the box then definitely you know where is the vanilla ice cream and you can just take it out.

In that case, the probability of getting a vanilla flavor ice cream will be 5/30. That means, there are 5 ice creams of vanilla flavor and they are to be chosen from 30 ice creams. 30 ice creams are for all the 6 flavors. So, this probability is simply 5 by 30 that is the basic elementary definition. And now, you know, what is the interpretation of 5 by 30. And if you try to take out this ice cream for a large number of times, and then possibly that this probability will come out come out to be 30, that I do not need to explain you now, as soon as I say the probability is this you will understand that what is the real interpretation of this value.

And then we try to find out the probability of the complimentary event of A that is the probability of not finding a vanilla flavor ice cream from the box. So, that is simply the complimentary event. So that is going to be 1-P(A). So, P(A) that we have obtained is 5 by 30. So, this will become 1-5 by 30, which is 25 by 30. As simple as that.

So, you can see here that how the computation of such probabilities is not difficult. And many times, you will see that when you want to compute a specific type of probability, then computing that probability through the complementary event sometime become easier. So, that is why I am was trying to give you this rule and this example.

(Refer Slide Time: 08:23)

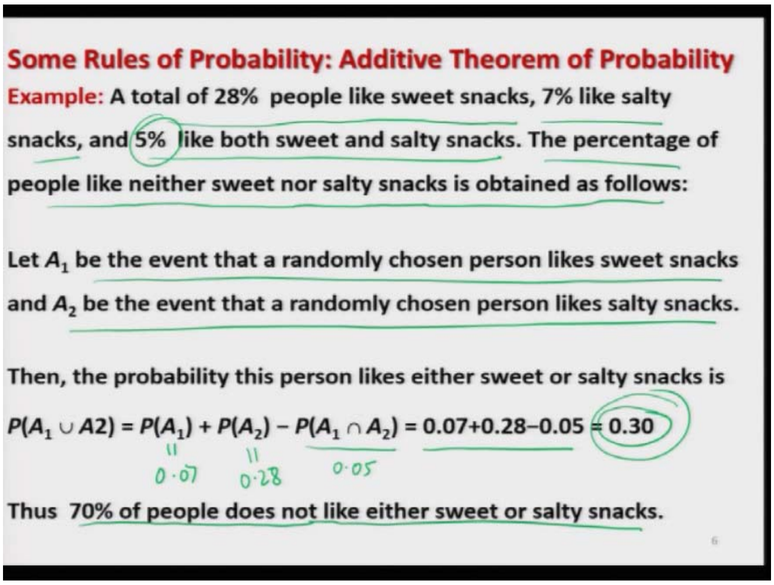


Now, I try to show you here one more theorem, this says that let A_1 and A_2 be two events which are not necessarily disjoint events. And if you remember, we had discussed this type of results, when these events are disjoint. We had understood the general result $P(A_1 \cup A_2 \cup \ldots \cup A_n)$ or so on. But now here, I am trying to say that these two events are not disjoints that means they can occur together there will be some common part between A and between the two events A_1 and A_2.

In this case, the $P(A_1 \cup A_2)$ that is the probability of occurrence of A_1 or A_2 is given by $P(A_1) + P(A_2) - P(A_1 \cap A_2)$. And definitely in case if even A_1 and A_2 are disjoint, then $A_1 \cap A_2$ becomes ϕ and the probability of this thing becomes 0. And we get here only $P(A_1) + P(A_2)$, that we already have done.

So, we know that the meaning of this "or" in the that "or" the probability of occurrence of A_1 or A_2. This in statistical sense is either A_1 is occurring A_2 is occurring or both of them are occurring. This result can be extended to more than two events also, but firstly, let me try to take an example to show you that how the things are working.

(Refer Slide Time: 10:02)

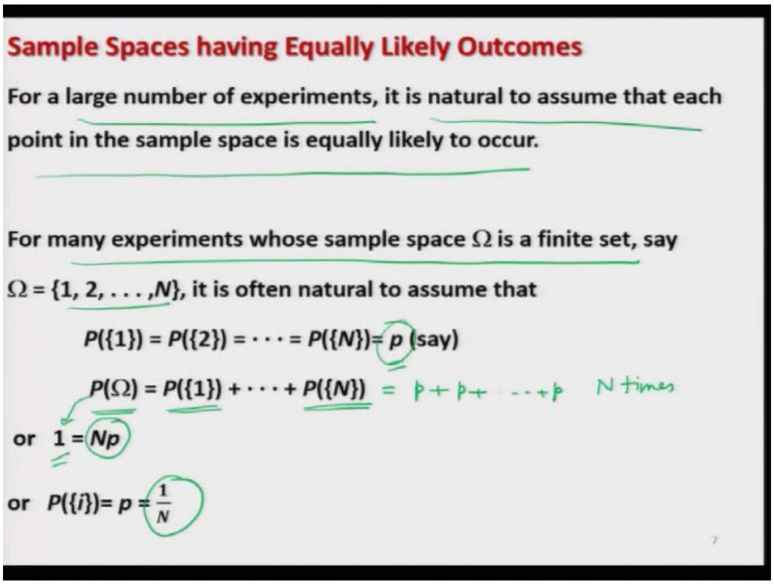


Suppose a total of 28 precent people like sweet snacks, 7 precent like salty snacks, 5 precent like both that is sweet and salty snacks both. Now, we want to know the percentage of people who neither like sweet nor salty snacks. And in order to find out such probability we have to use this rule $P(A_1 \cup A_2)$ because A_1 and A_2 are not here disjoint because there are 5 precent people who are liking sweet and salty snacks both.

So, in order to solve such problem, we try to define two events and this you are going to decide on the basis of options that you have. For example, you have here two scenarios, one for sweet snacks and another for salty snacks. So, we would like to have two events here say let A_1 and A_2 be those two events such that A_1 be the event that randomly chosen person like sweet snacks and A_2 be the event that the randomly chosen person likes salty snacks.

Now, you know the $P(A_1)$ is 0.07 because this is 7 precent and $P(A_2)$ that is here 0.28 and $P(A_1 \cap A_2)$ is simply here 0.05 which is 5 precent. And if you try to write down this probability you will get here 0.30. So, now, you can see here that there are actually 70 precent of the people who does not like either sweet or salty snacks, because there are 30 precent people who are liking actually either sweet or salty snacks. So, remaining will not like it.

(Refer Slide Time: 12:28)



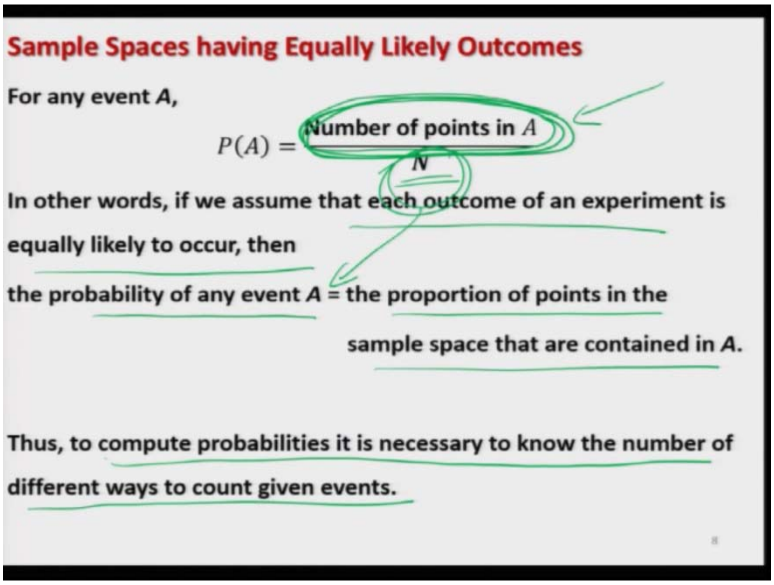
Now, let me try to give you here one more rule, you see, whenever we are trying to conduct an experiment for large number of times, then in those situations, from the practicality point of view, it is very natural to assume that each point in the sample space is equally likely to occur. In such cases, how are we going to compute the probability of an event?

So, for many experiments, whose sample space Ω is a finite set that means there are finite number of values. Suppose there are some numbers like 1, 2, 3, 4 up to N, these are different events, event number 1, event number 2, event number N. And suppose all these events are equally likely, that means the probability of occurrence of 1, 2 or say N, that is the same.

Suppose this is equal to here the small p like this. Now, we know that probability of Ω that will always be equal to 1, will be equal to probability of occurrence of event 1 plus event 2 plus probability of occurrence of event N and since, we are going to assume that each of this value is P, so this will become P + P + P up to P, how many times? N times.

So, this value will come out to be NP and this probability of Ω is equal to 1. So, probability of gaining any of that event I out of 1 to N will simply become 1/N and that is what we are trying to do in many situations. For example, when we conducted the simulation through the R software for obtaining the number of heads, tails or obtaining the number of the 1, 2 3, 4, 5, 6 for the roll of a dice, we had assumed that the probability of occurrence of any of the number or head and tail when we are trying to repeat it for a large number of times, they are the same. So, this is what we want to have.

(Refer Slide Time: 14:43)



Now, I try to once again come back to a basic definition and after that I will stop the lecture. Now, my objective is this after understanding this basic topics, basic definition fundamental definitions, you have now understood what is really probability. And what is the real interpretation of this thing.

You have understood that in case if you are trying to compute the probability in terms of relative frequency, then what is the interpretation of that value, but ultimately you are interested in the computation of such probabilities. And whenever you are trying to compute such probabilities, you are always trying to find out the number of points in the event A divided by total number of times the experiment is repeated.

So, now in case if you assume that each outcome of an experiment is equally likely to occur, then in that case we know that the probability of any event A is the proportion of points in the sample space that are contained in A. And this will essentially give you the relative frequency. So, now, when you want to compute this probability, did not you think that you need to know this value in the numerator number of points in A. And this is what you have to actually find. And this is what exactly you were doing in your elementary classes.

For example, you have done many examples where you will say that okay if a box has 5 black and say 6 blue colored ball and if you want to find out 2 black color balls and 3 blue color balls, what is the probability? You use to make different types of combinations and you will try to find out that what are the total number of ways in which the event can can occur and this is exactly the same thing you are trying to find out the number of points in A. They are also saying the same thing that the total number of ways ways in which the event can really occur.

So, now, once you can compute this number of points in A then N you know, and you can very easily compute this probability of an event A. Thus, to compute the probabilities it is necessary to know the number of different ways to count the given events. Now, when you have understood that in order to compute a probability, you need to know the total number of favorable events or the total number of ways in which the event can happen, we need to know that how to count these things. And it is not only counting, you have to know that how to know those values based on our software.

For example, you have done different types of things in your earlier classes like combinations, factorial, etc. What are those things? They are simply different ways of counting. So, now if you can understand that how to count the total number of possible events, you simply have to divide it by the total number of events, that is not difficult.

So, now my next target after giving you these basic fundamental rules and very simple example, my next target is to explain you that how you can count different types of combinations, permutations etc etc, and how you can use them in the software. So, now I stop in this lecture, I will simply request you that you please try to go through with these rules. Try to think about them, try to set them in your mind and try to see how are they working. What are they trying to give you and what type of information, how that information is being transmitted to you.

As we have discussed in the beginning itself, that is statistics is the language of data, now you have developed one tool probability theory which is giving you some number between 0 and 1. Now, this is your job to understand what is this number between 0 and 1 is trying to inform you.

So, now in the next lecture, I will take some basic methodologies for counting, but I will request you that you try to have a quick look on such concepts which you have studied in your classes that will help you in the next lecture. So, you try to revise it and try to have a quick look on that on different number of ways to count and I will see you in the next lecture with those topics. Till then, goodbye.