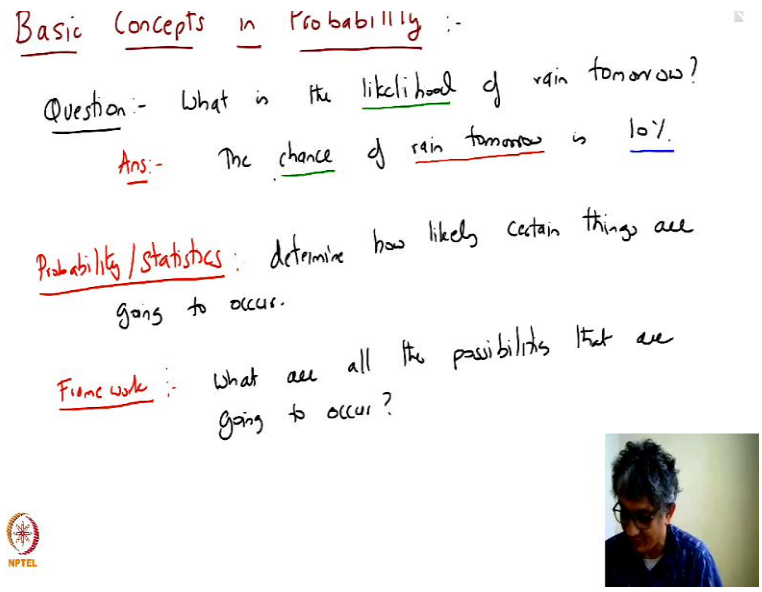
Introduction to Probability- With Examples Using R
Professor Siva Athreya
Theoretical Statistics and Mathematics Division
Indian Statistical Institute, Bangalore
Lecture - 01
Sample Space, Events and Probability

(Refer Slide Time: 00:17)



So, welcome to the first class on probability, that I am offering this semester. So, I would like to begin with some basic concepts and problems. So, before I start setting up the structure; let me post some questions. The first question should perhaps deal with daily life. So, for example, let me write the question down. So, what is the likelihood of rain tomorrow? Or pertinently, you might look at another bulletin and it might give you an answer, that states that the chance of rain is 10 percent. The chance of rain tomorrow is 10 percent.

So, I would like to focus on this, on both the question and the answer. One is that, in our heads, we sort of have an understanding of what, when we say likelihood or chance, what it actually means to us. Some may use the word probability, some may use likelihood, some may use chance; but we use these words interchangeably. That is one thing, that I would like to sort of focus on in these 2 sentences.

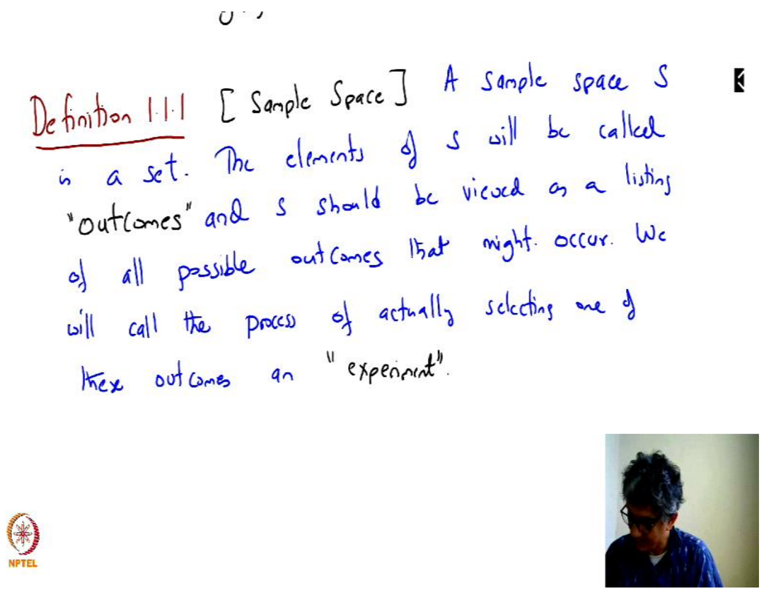
The next thing is the answer of course, that the answer is 10 percent. So, this number means something to all of us. So, if you ask somebody on in the street, what does this sentence mean, they might give you an answer that differs from what you think the correct answer is. And of course, the other thing you like to focus on is, what we have done is the event of interest. So, the event of interest is rain tomorrow. So, there are 3 things in this 2 the question and answer. One is the likelihood or the chance. One is the event of interest, that is rain tomorrow, and one is the number, that is 10 percent.

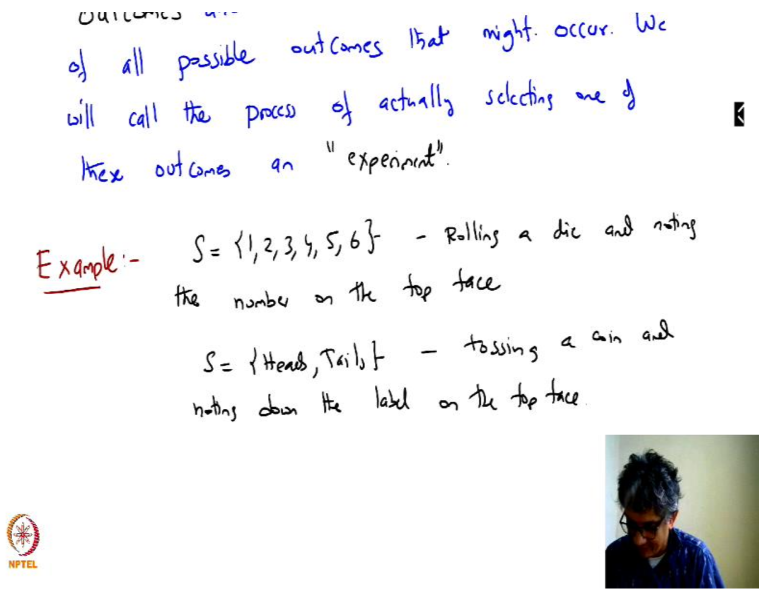
So, this is a feature of probability and statistics, the feature of probability and statistics. So, both these subjects always want to determine, how likely the certain things are going to occur. So, they always try to determine how likely certain things are going to occur. So, one might again determine that this statement is little to vague, but you should keep it in the context of the question that has stated before. There are certain events that you have, and you try and understand how likelihood are. So, both these subjects deal differently. We will try and understand the probability part of it in this course, not the statistics part.

So, for this, to do this, I like, I said before, I must understand what I mean by likely. What are certain things going to occur? And before I begin, the whole structure of understanding these concepts; the first is I like to build a framework for doing the above. And the first question that we would like to answer, answer is what are all the possibilities, before we determine what is likely and not likely; we need to know what are all the possibilities.

Once we know that, then we can set up how likely is it. Something that we are interested in is going to occur. So that is what, I meant by saying in the first question that, if you ask somebody on the street that, what does it mean to say chance of rain tomorrow 10 percent. The person must know, what are all the likely possibilities, and then what is likelihood of the given event of interest.

(Refer Slide Time: 04:42)





So, here is the first definition. So, I will call it as definition. Definition 1.1.1. So here, I would define a sample space. So, sample space S is just a set. So, by definition this is just a set. A sample space S is a set. We refer to the elements of S as outcomes. So, the elements of S will be called outcomes. So, each element in the set is an outcome, and S should be viewed as a listing of all possible outcomes, and S should be viewed as a listing of all possible outcomes, that might occur.

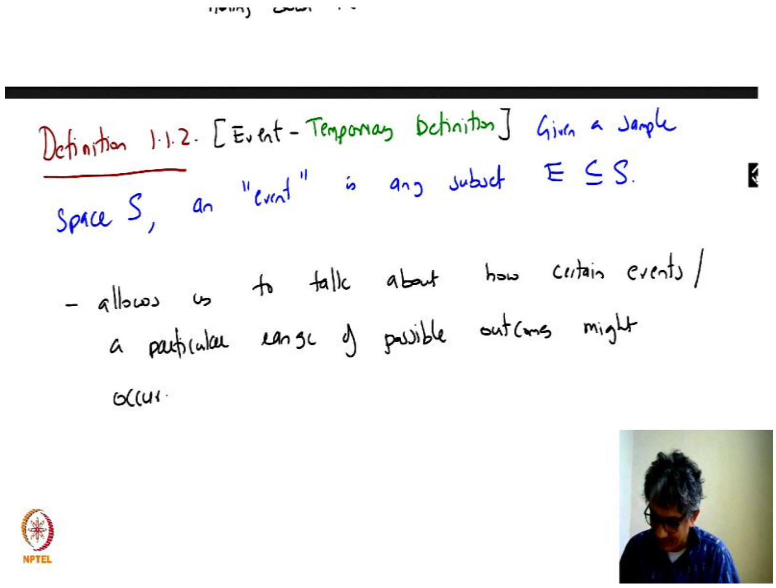
So then, once you have this, what I will do is; I have a listing of all possible outcomes. Then I will call, we will call, or we will name it, how we want to phrase it, we will call the process of actually selecting one of these outcomes as an experiment, an experiment. Let me put an inputs. We will put (())(7:13), we will put experiment inputs. So, on my view experiment as conduct experimental outcome happens, or you can view as the process of actually selecting one outcome as an experiment. So, as the sample space.

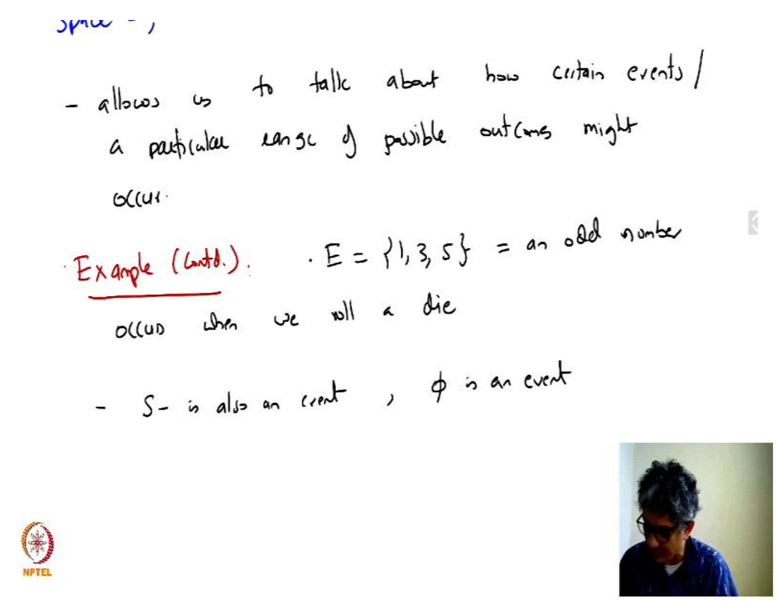
Now, let us give a quick example. So, let us say I will take the S as 1, 2, 3, 4, 5, and 6 as my set S as sample space. Then I would this lists all possible outcomes of a rolling of a die. So, I would think of my experiment as rolling a die. I am rolling a die, and noting down the outcomes on the top, noting down the number on the top. So, in some sense, S is all possible outcome that will happen in this experiment, and the experiment is rolling a die, and noting down the number on the top face.

Or you could think of it as S is just, let us say heads or tails. And you could think of the experiment in associating this as tossing a coin, and noting down the label on the top face.

Now, once we have a list of all possible outcomes, we now should understand what thing we are interested in.

(Refer Slide Time: 09:28)





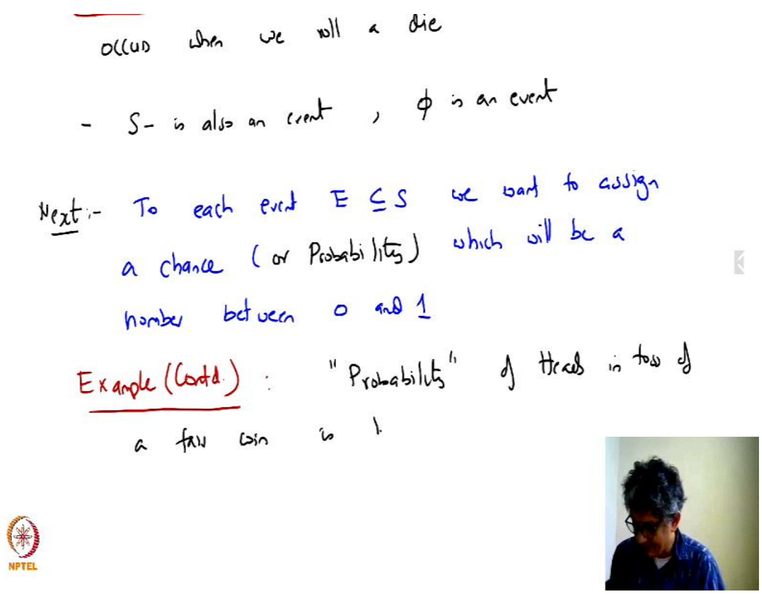
So that is the next definition, I would like to start here. So, this is called an event. Definition 1.1.2. It is called an event. I will just alert you little bit. So, I will give a temporary definition, not because I do not want to give you the full definition at this point in time. It is just that, the precise definition will take some time to develop. So, for our purposes, now we will just assume, given a sample space S, an event is any subset E of S, any subset E of S is an event.

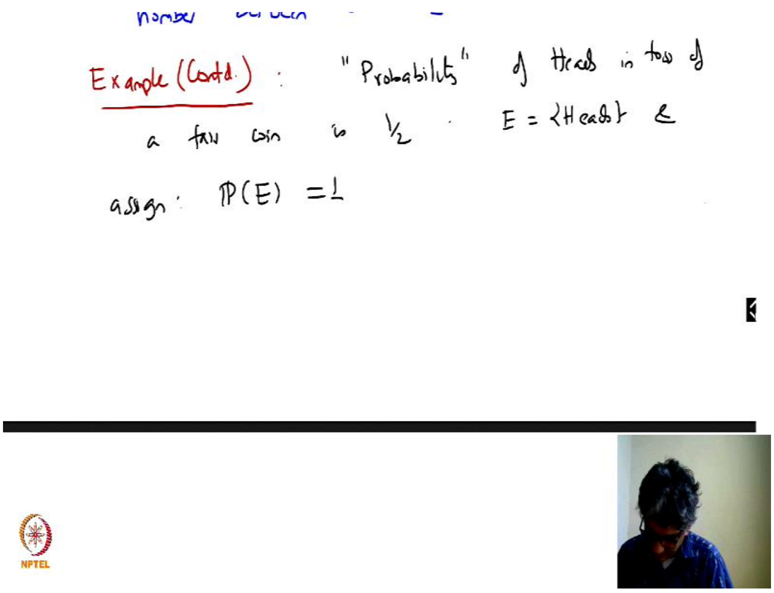
So, what does this allow us to do. This allows us to talk about how likely is it a possible a subset of outcomes can occur. So, the definition allows us, so allows us to talk about how certain interests, certain events of interest, or a particular range of possible outcomes might occur. So, for examples let us go back to the example that we had before. In this example, we roll a die S of E. So, in our example let me come to let me say example continue. You could say an odd number occur. You could take the event E as 1, 3, 5; as the, an odd number occur. Odd number occurs when we roll a die. That is that I think.

So, the other remark, that I would like to sort of make before I move on to the next event is that S is also an event, is also an event and is the entire sample space in there. This means that anything can occur. The other thing is that the empty set is also an event, that means nothing occur. That is what I mean by all possibilities and no possibilities occur, that is what the empty set says.

Now, comes our final task. So, you have already understood now what all can occur, and we have given a name for the event that will describe how certain things of interest that you want to be defined as. Now our last part in this whole structure is to understand what it means to assign or state how likely something is, the last one.

(Refer Slide Time: 13:07)



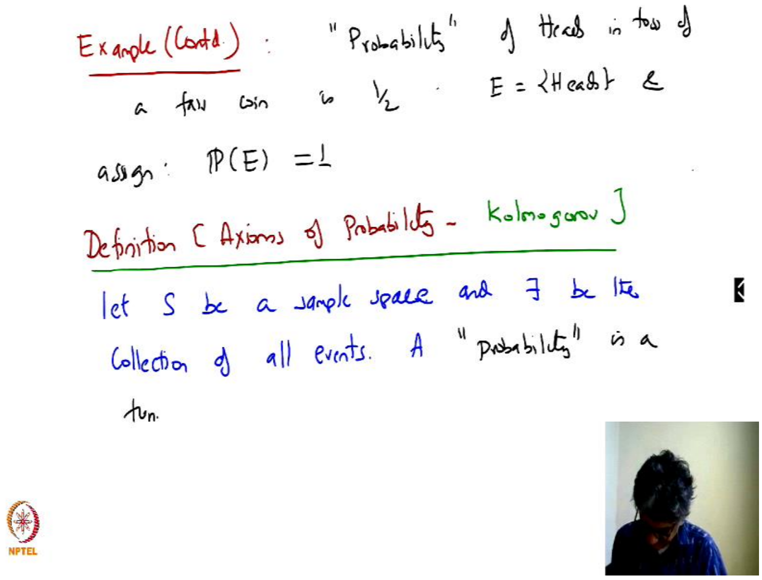


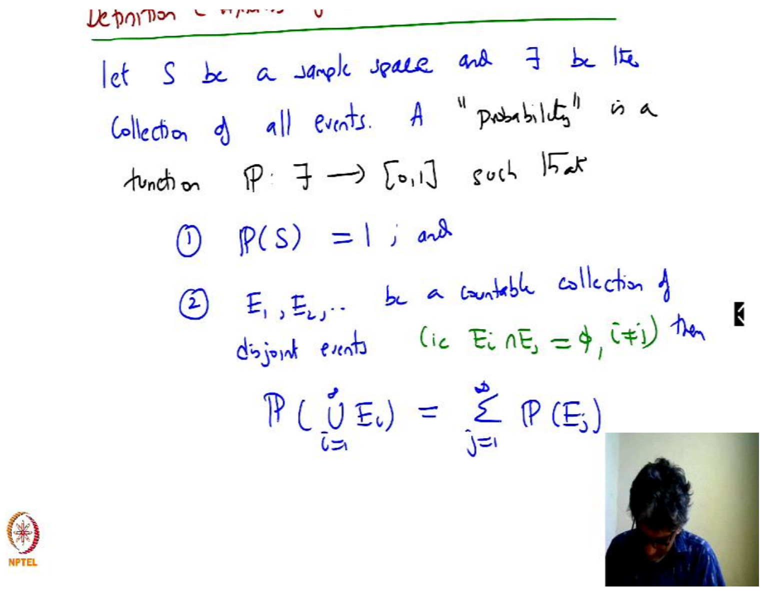
So, what we will do next is, so next what we will do is, to each event, each event E subset of S, what we will do is, we will assign, we want to assign a likelihood, or how likely it is a chance, the chance of rain tomorrow, how likely. Or what I will call now as probability, which will be a number between 0 and 1.

So, now when the experiment was performed, a particular event has a certain chance of occurring. So, let us go back to our example again, so what we might say is that, example. I might say the following that the, I might say this, even though I have not defined it yet. The probability of heads in the toss of a fair coin is half. That means all I have done is that, I took the event E to be H, that is heads. That is heads happens, and I assign the probability of E to be equal to half, that is what happened.

But one has to define this sort of setup, you cannot go to every event and every experiment and do it separately. One has to have a formal definition of what we mean by a problem. So, these are given by a set of axioms, which I also write down now as the last concluding part of the framework.

(Refer Slide Time: 15:15)



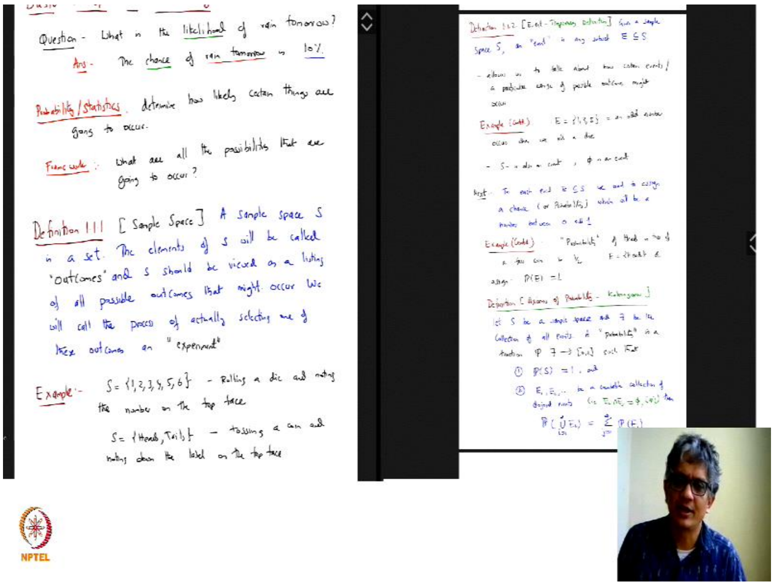


So, let me add the definition. So, definition. So, these are called the axioms of probability. And even though these were around for quite some time, this formal notation that are going to use was given by a famous probabilist called Kolmogorov, which is what we will describe. So now, here I will start with 2 notations that we are already familiar with. Let S be a sample space, that we have, that is given to us.

And let us script F be the collection of all events. So, essentially all subsets of S I will call F. So, then I will define probability is a function P from this event space to 0 1, such that, what happen, 1 is the probability of the whole sample space is 1; that means all possible outcomes are listed in S, that means if you perform an experiment, something in S will occur. And another usual important property is that, if I take, if I split my sample space into disjoint events or if I take a disjoint collection of events, we might have a countable collection of events, of disjoints events.

So, what I mean by disjoint events, I just mean that, that is E i intersection E j is empty for i not equal to j, that means there is no intersection between 2 of them. Then the probability sums up. In the sense that probability of the countable union of the event that is any, anything in one of the events can occur. Because in disjoint that is the same as the sum from j equal to 1 to infinity the probability of E j.

(Refer Slide Time: 18:16)



So, let now just quickly describe everything in one summary, so we all are on the same page. Where is my lecture 1? Right here. Let me make it smaller. Move it up. Move it down. Make it smaller. Very nice. Make it smaller may be. So here, on the left-hand side, I start off with this idea of what I thought was the key question to understand. So, all of us are interested in how likely a certain event is occurring.

Then I said, to do that I need to know first what all can occur? So, I define what is called a sample space. So, sample space S, like you noticed here, is just the set of all possible outcomes. So, we call outcomes and S should be viewed as a listing of all possible outcomes, and you think of an experiment as choosing one outcome in S.

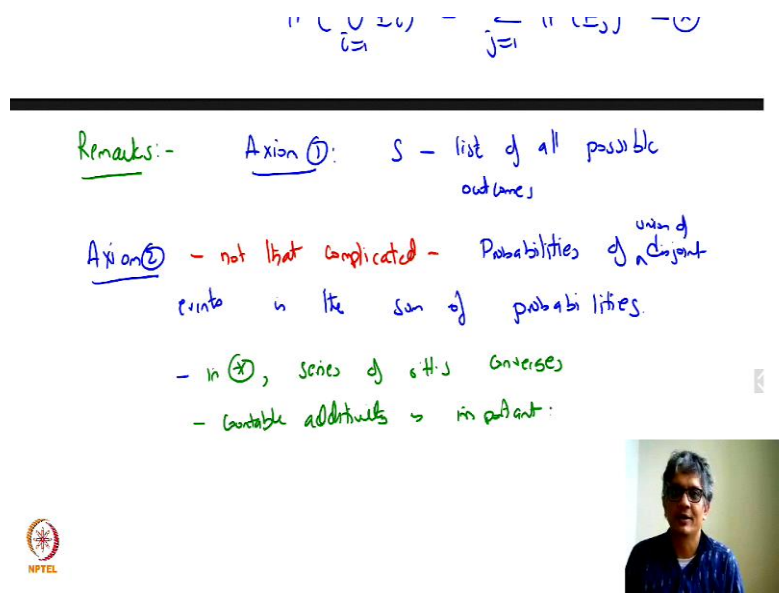
So, a quick example that we discussed were, one was this example of rolling a die, the other was heads or tails in tossing a coin, and so and so. Once we define sample space, on the right-hand side I began by defining what an event was. An event as of now for us will be any possible subset of the sample space. That is an event. It is a temporary definition because event will have to have some structure. We will soon learn later on that we cannot assign probabilities to all subsets. We will learn that later on. But as of now, we will think of any subsets.

So, example we could take the rolling of a die example right here. We could take E as 1,3, and 5. That is an odd number will occur. One thing to notice, S is also an event, so is the empty set. Next, we did was, we said we had assigned probabilities to everywhere, that is a number between 0 and 1. To do this, I gave intrude example, if you toss a fair coin, you will expect heads to come half the number of times. That is the probability of heads is half. That means the probability of E is half, that is E is defined as heads. But I do not want to do that. I want to go and do it in a systematic way, in one shot for all possible cases.

So, I defined S to be my sample space, and take F to be the equation of all events. Then the probability is the function P from F to 0 1, such that P of S is equal to 1, and it has to have one more axiom that is, if you have countable collection of events which are disjoint, then their probability is sum of them, that is the probability of their union is the sum of their probabilities.

Like I said before, the first axiom is straight forward. It just says that, S indeed is a listing of all possible outcomes, that was P of S is equal to 1. The second is not as complicated as it look. It just says that, when we have, when we combine disjoint events, the probabilities sums up. Sorry, let me now go back to the hundred percent view.

(Refer Slide Time: 21:45)



let S be a sample space and $\mathcal{F}$ be the collection of all events. A "Probability" is a function $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that

1. $\mathbb{P}(S) = 1$; and
2. $E_1, E_2, \ldots$ be a countable collection of disjoint events (ie $E_i \cap E_j = \phi$, $i \neq j$) then

$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{j=1}^{\infty} \mathbb{P}(E_j) \quad -(*)$$

So, now let me just play a little bit about the 2 axioms, that I had said, these 2 remarks. Axiom 1 indeed says that, we think of S is the list of all possible outcomes. That is all that is saying. If you form an experiment, and S is the sample space, then one of the outcomes in S should occur. Axiom 2 is not so complicated, is not as complicated, is not that complicated, though it is written in a very complicated way. All it means is that disjoint additivity happens in probability. All it means is that probabilities of disjoint events, policy of union of disjoint event, is the sum of them.

Though I must stress that is the, there are several things to stress here. That is, in, so when I, let me call this as a star. In star, so there are several things to stress, one is in star. If you know calculus, and aware of the thing, the series on the right-hand side converges. That is what that means. Otherwise, if you do not know calculus that much, just think of it as just adding up countably many problems and giving you a finite number. And the countableness is important, the countable additivity is important, which we will stress later on. It is not enough to assume just finitely behaviour. But once you have countable additivity it also implies finite behaviour.

So, such things will not be, this sort of subtleties will come clear as we discuss more and more general probabilities spaces, but this is where we are. So, as of now we have defined a sample space, an event space. So, sample space, events, and a problem, and probability is just a function from the events to 0 1, such that the probability of S is equal to 1 and the countable disjoint events, the probability of countably is sum of them all.