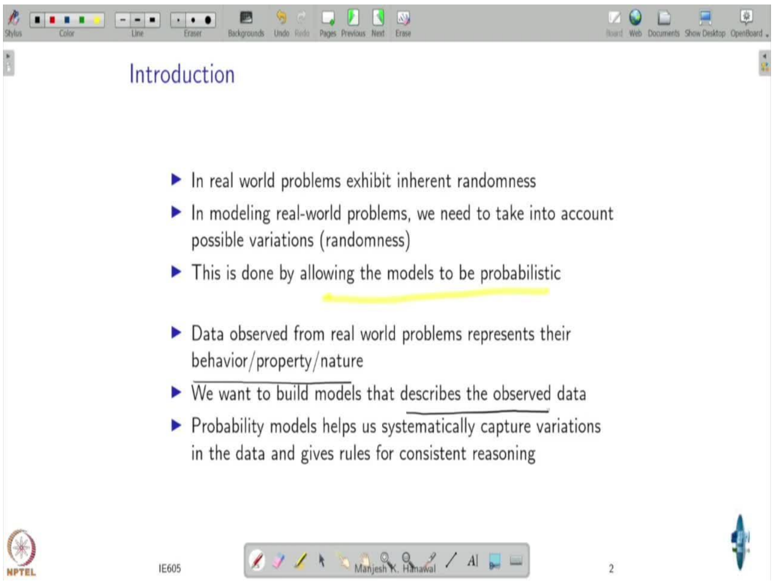
Engineering Statistics
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Lecture 01
Introduction to Probability

Okay, so we will get started with our first lecture here. So, what I am going to do is, get started with the introduction to probability. But in the statistics, I am expecting you people already know a little bit of probability. So, I will just quickly go through and those of you are especially M tech or IEOR students will be also doing IE 621 there you will see this more, but still I want to go through for the sake of other students in the class.

(Refer Slide Time: 00:51)



Now, as you all can imagine, there is nothing deterministic in life, majority of the things are like random there is certain uncertainty associated with this, like even they say that tomorrow sun will arise or not, that is going to be not sure sort guaranteed there is sub tense, maybe small, very, very small, but there is a possibility that it may not happen. So, there is randomness, in many real-world problems, we believe, and but we want to understand them.

So, that you should have to deal with those problems, we need to understand. And that is why we need to take into account the randomness, and now how we are going to do this? If you have to deal with these real-world problems, which you know, they are going to have certain uncertainty or randomness associated with this, then what we are going to do is in that case, if you have to analyze them, we are going to allow or build models, which themselves are probabilistic.

As we go along in this course, you will see that what are the probabilistic things that we are going to use in modeling real world problems? And also, what are the real-world problems that we are going to see, we actually do not know what is the probabilistic behavior or the probability model they are going to use, but what we get to know is their behavior, property or nature through their data that they generate.

So, for example, we know that weather. Weather, we cannot predict weather, it is difficult, it is a random quantity like I do not know, whether next year it is going to rain heavily in Mumbai or not or it is going to be the monsoon will be going to be good, bad or it is going to be normal. There is a certain kind of underlying randomness there I do not know according to what probability that is going to happen.

But anyway, next year, I see whether it was a normal monsoon by the like, by the end of when the monsoon season ends, I know that it was normal, or deficit or it was I mean over, or like it was more than normal. So, that is the data I observed. And that data is like what is generated from that underlying real-world problem.

Now, what are we trying to do? We get to observe this data, and observing this data is what we are trying to do and build the underlying model. Maybe like weather as a real-world phenomenon is happening according to some probabilistic model. I do not know that model behavior, the only way I see it is through the data generated so I try to see the data and go back and try to see how to build a model that can potentially better explain that real world problem.

And why probability models? I really do not know like, what is the model that is being used to decide all the weather conditions? But we know that there is some kind of randomness involved in that and that is why to at least understand that systematically, we want to have to start using some model and because of this randomness, we start thinking about some probability models here, and what we need to ensure is, if you have able to suppose let us say, there is an underlying model according to which the weather phenomena happens.

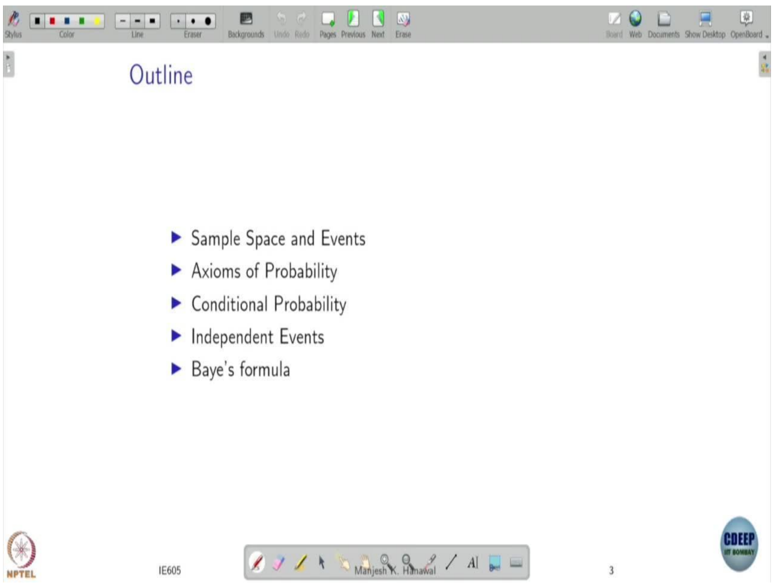
I suppose you know that weather phenomena are somehow able to model it well, then we are going to be sure about whether the weather model you captured is going to be good or bad. You are going to see whether it is good or bad based on the data that you are going to observe and see that that data being generated is consistent with your model. So, let us say that weather again is very classic example weather.

Let us say our metrological department builds a model, they build a model collecting our data, what happened in the last 10 years 15 years 20 years 50 years whatever. And now, they build a model and then they make a prediction for the next year, they may say that, maybe monsoon is going to be normal. And then next year, you actually see whether the monsoon is normal or not. If the monsoon happens to be normal, then maybe the underlying probability model you have built is actually close to what nature is trying to do. And you are good.

So, I\you have tried to understand that probabilistic model if not then maybe something like you have your model is not as good your probability model is not good, you maybe need to improve that and that is what happens like the weather prediction algorithm exactly does that like they make a predictions and if things are as per them fine, if not, they will go back take the data and try to improve the model. So, that is why it is the whole thing like you see that there is probability model and data here we need to have probability models, which tries to govern and try to model the real-world phenomena. And data is what we observe.

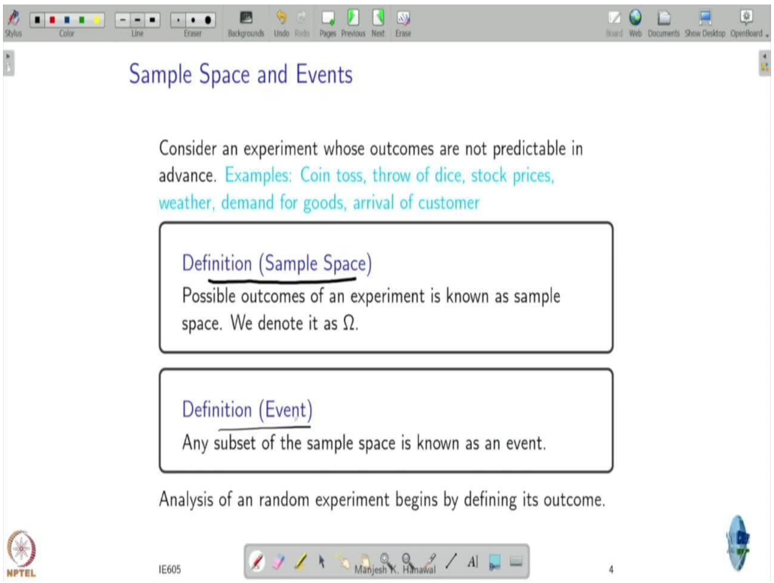
And see that whether the data that is being generated is consistent with our probability model, if it is so then we know that we have modeled the real-world phenomena well, and if we model the real-world phenomena, well, that is good for us. Like we can predict things properly and accordingly we can take our actions.

(Refer Slide Time: 07:26)



Now, let us start building the basics of probability. Maybe some of these things are already being parallelly done and in IE 621. But I will go through this quickly today, so what we will start talking about in this class is about simple sample space events, Axioms probability, conditional probability, independent events and Bayes' formula like let us start to finish this today.

(Refer Slide Time: 07:52)

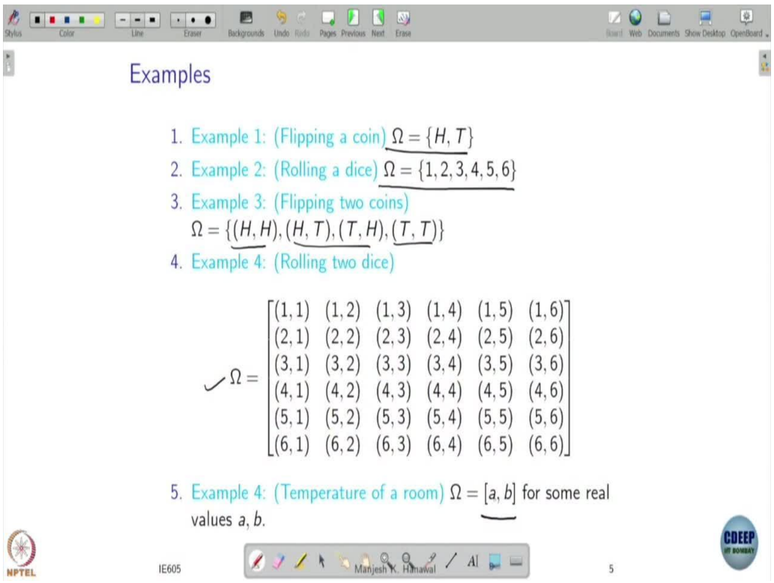


These are all classics, which everybody who knows basic probability already knows. So, if you are going to consider any experiment now and then, we are talking about random experiments. Like, just think of like, let us think of weather itself, like weather is like an experiment like you want to see what is going to happen tomorrow, that means whatever the underlying conditions that are going to make something to happen, maybe like, make the weather behave in a certain pattern. So, these are all like, we are just going to treat it as experiments, random experiments, the outcome, what we are going to see they are coming through this random experience.

Now, to understand that we will not go into anything complex, like weather, we do not know whether there are a lot of things in weather. Like there is temperature, humidity, and what else maybe perception, that density, a lot of things are involved, that all are going to govern how the weather behaves. But to begin with, we will not get into that complexity, we will strip down the things to the bare minimum possible. So, what is the bare minimum possible we will go to the case of simple coin toss or a throw of a dice to understand all these things, these are still random quantities, but they are pretty simple to understand and reason.

So, we know about coin tosses. This coin toss is a random experiment or it is a deterministic experiment. It is going to be a random experiment. Because you cannot apriori predict what is the thing you are going to say whether it is heads or tails. So, now, the first thing we are going to look into is sample space, we are going to say possible outcome of an experiment as a sample space and we denoted by omega and the next thing of interest for us is something called event and we are going to treat any subset of the sample space is known as an event. If you want to understand a set up of a random experiment or understand or analyze that random experiment, the first thing you need to do is have some understanding about what are these possible outcomes and what are these possible events.

(Refer Slide Time: 10:50)

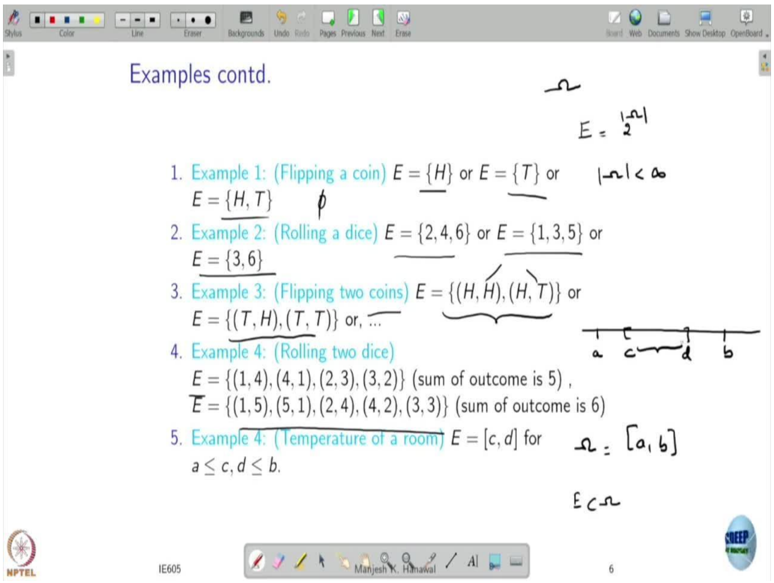


Now, let us look into that. So, in the case of coin flipping, we know that the outcome is going to be either head or tail. So, that is why the sample space is simply going to be head or tail H represents head T represents tail and in terms of rolling of a dice the dice will have six faces. And any one of them comes. I do not know which one and that is why they are represented 1, 2, 3, 4, 5, 6 and other simple examples like instead of one coin you may want to flip two coins and, in that case, there are four possibilities. Either both of them may show head both of them may show tail or each of them show different things: first one shows head, second one tail or vice versa.

And similarly, if you are going to take two dice there are now 36 possibilities. And that 36 possibilities are written in this matrix and that will constitute your Ω . And now your outcome or a sample space can actually need not be as simple as that it could be actually continuous space or interval.

For example, if you are interested in room temperature and that room temperature the possible range of that temperature could be anywhere between some value a and b for example, if you are interested in room temperature in Mumbai it could be something anywhere between 20 to 45 degrees So, then in that case, we are going to say my outcome sample space is the interval [a, b] and any value in the interval [a, b] can happen.

(Refer Slide Time: 12:53)



Now, events. We wanted to talk about events. Now, we said that the event is nothing but a subset of sample space. So, if you take the coin toss, subsets are H and T or both HT and in fact null set can also be there. Now, what does the event that H means? That means like you are interested in the event heads and T means we are interested in the tail and HT means we are interested in either of them like whatever comes fine and that is here.

And what is phi here nothing happening. Nothing is happening. But when you are going to toss a coin, Can nothing happen?. No right either head tail happens or if you are interested in both either is fine but nothing but if you feel null set still trivial event we can consider and in the rolling of the dice you may be interested in knowing an event in which the outcome is divisible by 2 or even number in that case your event is 2, 4, 6 or you may be interested in an event where your outcome is odd number in which you will be interested in 1, 3, 5 or you may be interested in an event where your outcome is divisible by 3 in which case every event is going to be 3, 6 like that. And now if I have omega, as my sample space, how many events will be there?

Student: $2^{|\Omega|}$

Professor: Two to the power?

Student: cardinality in omega

Professor: And it includes a null set?

Student: Yes

Professor: So, it will be total to the power cardinality of Ω will be your thing, but this is makes sense when you are omega is finite, that is you are interested our possible outcomes are finite, like if not, this could be just infinity, but that is fine. Any event, which is a subset of my outcome, I am going to treat it as an event. In terms of flipping of coins, like, can somebody tell what this event represents?

Student: (())(15:43)

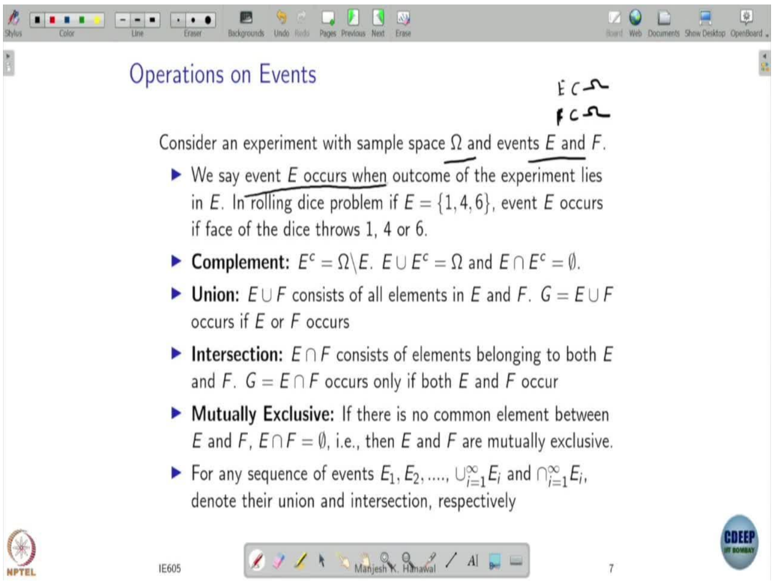
Professor: Yeah, here it is. Yeah, here it is saying that, I am interested in the case where the first outcome is head. So, in that case, the second one comes either head or tail, I do not care. And similarly, this event is representing that my first outcome should be tail. In the roll of dice, again, you may be interested in an event where the sum of the two faces is going to be 5. So, in that case, these are all the possible outcomes we will be interested in. You can consider all these possible outcomes as your events.

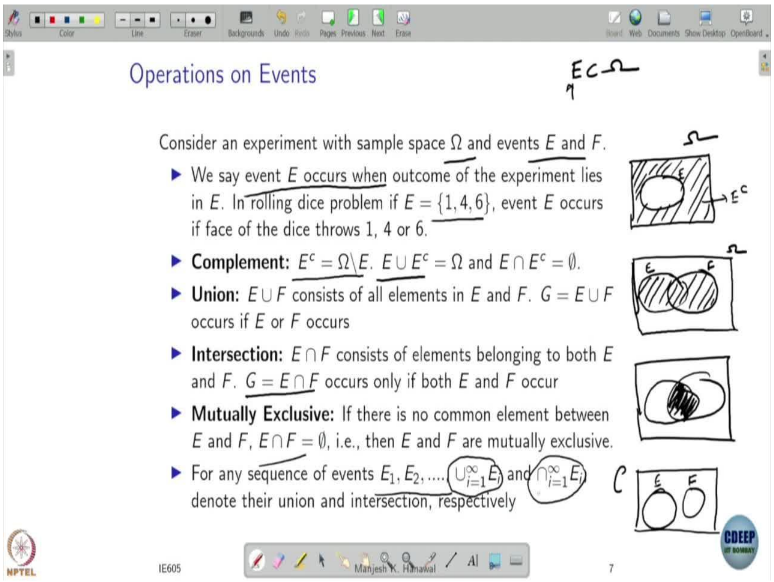
Now, in case of a temperature, we say that, let us say [a, b] was your interval, your possible outcome, this was our Ω and event can be any subset of this. Let us say like let us say if my this is my a, and this is my b room temperature. And I will take some interval here between this c and d. And then in that case, this is a subset and that could be an event.

So, instead of asking if your possible room temperature is between 20 to 45, you want to see that it will be cold in my room like that, in that case, suppose let us say the temperature will be between 20 to 24. And you will be interested in that particular subset now. Now, given that we have the sample space and events, you may want to do certain operations on events for example, like so, what is the chance that like, if you are throwing a dice that my outcome is even and also it is divisible by 3?

Now, you are trying to play with you are interested in two events now, like one event is an even number and another is divisible by 3 and you are not then happens, you are fine. And now, you want to know how to do it? Now you are trying to look into multiple events, and you are trying to do operations on that. So, what operations are possible.

(Refer Slide Time: 18:17)





So, let us take two events, which are events in omega. So, obviously, where this and F is also a subset of this. Now, let us try to understand when we say event E occurs, so we know that let us take I know E is a subset of omega. Now, I am talking about an event E happened, what does that mean? So, what it means is, let us say, so before that, let us say whenever any element in E is the outcome, then I am going to say event E has happened. So, what does that mean?

Suppose let us say you are rolling a dice and you are interested in the event 1, 4, 6, the face showing 1, 4, 6. Then we are going to say that event E has happened if either face shows 1 or 4 or 6. So, only if the face shows something like a 2, 3 or 5 then I will say that event E has not occurred. Otherwise I will just say that event E has happened.

Now, if I have to I am asking if event E has happened. Then I may be also interested in some events not happening, like I may want to, my room temperature is not between 40 to 45, so then what is that complement a complement is like a of an event E is nothing but you remove E from your omega and whatever remains is complement. For example, if you have this, let us say this is your Ω and you have some set E here and now all the things that are outside they are going to form an E.

Now, this naturally brings out a couple of properties on an event and its complement. So, if we are going to take union of E and E complement that has to be naturally omega. That is by definition and if you are going to take the intersection of E and E complement, that is null set because there is nothing common between them.

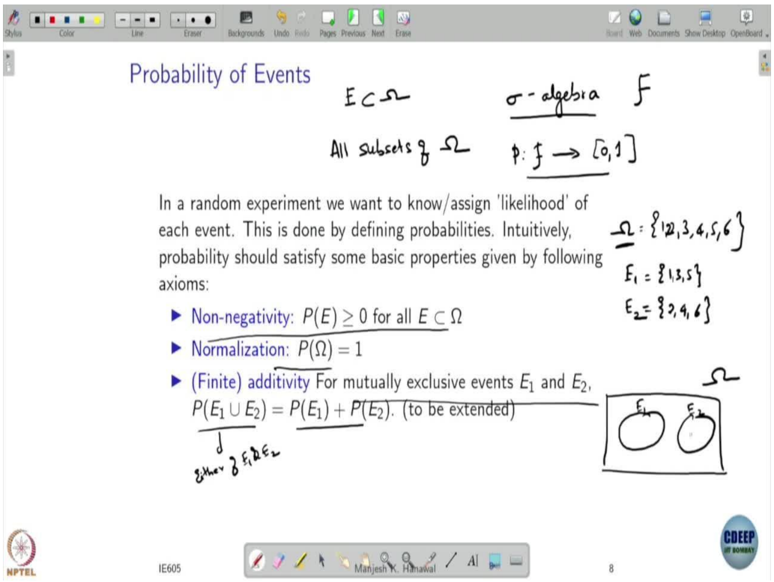
But let us this is a simple case when we looked into E and E complement their union and intersection that is clear because we know what happens but now let us take two things two sets and try to define their union and intersection let us say I have one set here and another set of F here, this is the omega. What is the union? The union is simply nothing but all the elements in this. So, which is all of this and intersection similarly we are going to define to be simply all the elements which are common in that.

And now let us say G is an intersection of E and F, two events. Now, what, like G is a new event I have defined using event E and F. Now, when I am going to say event G happened, it should mean that both E and F happened. That means the element which is common in both E and F, if that has come as an outcome, then I am going to say that event G has happened.

That means basically the outcome should be from the intersection, if this has happened, then I am going to say that event G has happened. Next mutually exclusive. If there is nothing common in between these two elements. Let us take two events E and F. Now, what we are saying is, if you take the intersection if it is null, that means there is nothing common between them like in this case. E and F there is nothing common between them. And in this case, we are going to say the events as mutually exclusive events.

And now, so, this is a simple case of two events we have consider, but the same definition of union intersections applies even if we have countably many events, for example, let us say E1, E2, all the way up to ∞, then they are union means that means this is if an element belongs to any other event, it is there in this union and similarly, an element belongs to the intersection if and if only if it is belongs to each of these events. And even though I have written it for this countably many events here, this applies even if it is a finite, I mean, this is a standard definition. So, all of you understand the difference between finite and countable here. Anybody who does not understand it? Okay fine.

(Refer Slide Time: 25:04)



So, now let us get into probability. So, you people are already in IE 621 you talked about σ-algebra.

Student: (())(25:14)

Professor: So, you learn it but I am not going to go into that. So, are any of you aware of what σ-algebra is? Probability space: what is probability space? Just tell me, let me let me just make it simpler for you. I just want to know, do you know σ-algebra? One lets us just think like we talk about events.

We said events which are nothing but subsets of your omega now, let us collect all subsets of omega, we know that all subset overall there are $2^{|\Omega|}$ that size and now, so, there is a formal definition of σ-algebra I mean in this will not go into that you will anyway learn it in the other course and that is also not of so importance for us.

What we want to do is for every subset we would like to know the likelihood of that subset happening, let me ask this question like, if you have a simple dice you are throwing it what is the likelihood that the outcome is divisible by 2?

Student: 1 by 2

Professor: 1 by 2? What is the likelihood that that value is divisible by sorry, it is divisible by 3.

Student: 1 by 3.

Professor: 1 by 3? Now, why are you saying that? It is based on the likelihood. Because I mean everybody right now you are also assuming that it is a fair dice. So, each one of them showing up has the same likelihood, like now I am interested in particular like when I say divisible by 3 out of 6 possibilities are interested in 2 possibilities. 3 and 6, so 2 by 6 is how you are computing it.

Now, instead of going that, see a problem in probability we have to like define these things formally. Now, what we are now going to do is probability, we are now going to assign probability to each of the possible outcomes, each of the events. So, ultimately, when you are going to do an experiment, we will be interested in different events happening and you would like to know their probabilities or their likelihood or at least you want to model it you want to assign some probabilities or likelihood to them and that is where the probability comes into picture.

So, probability is nothing but like I mean, the σ-algebra is denoted by F script here, which is a subset of all. So, this is now going to be a function from the sub subset of all sample spaces to [0, 1]. And now, this the way we want to assign probabilities, we want to make sure that they satisfy certain basic properties and what properties they should satisfy? We are going to make some assumptions on that and that will make us make some axioms.

The first property we are going to assume is that the probability of any event is going to be non-negative. So, when we want to deal with probability, which when we are also going to talk about likelihood, we want some for non-negative numbers, they are like saying that the likelihood of something happening is -0.05 is not so intuitive. Like you want to assign some non-zero numbers, and that is why we want everything to be positive.

And also, when I consider the whole sample space itself I want that to be assigned some number, and that number should be the largest. Like for example like when I am saying let us say dice, if I tell any number is fine to me. The likelihood of that should be larger than any other possible event. So, I am saying let us say you are taking omega equals to 1, 2, 3, 4, 5, and 6. So, this is your dice which is the sample space.

Now, let us take two events, E1 and E2, E1 is let us say 1, 3, 5 that is odd number. And E2 is 2, 4, 6. And now, Ω is also an event, which is like entire thing, whose probability whose likely should who should be more E1 or Ω?

Student: Ω

Professor: E2 and Ω? So, Ω will have the largest likelihood. Whether it is covering basically, and that value, I want to normalize, I do not want it to take any value. And that is why I will put a normalization, I say that that value should be equal to 1. And last point, we assume that finite additivity, we say that if I going to take two mutually exclusive events, if I look into the union, so union of two events is going to be another event, the probability of that event should nothing but should be equals to the sum of the two events. So, for example, like let us take I have this omega, and I have these two mutually exclusive events, here, I am denoted as E1 and E2.

Then the likelihood of happening either of these is E1 U E2. That is either of, there should be nothing but then the probability that then the likelihood of these should add up because they are just separate. And, that is what we call mutually exclusive. And if they are mutually exclusive, then the probability of their union should be equal to the sum of their properties.