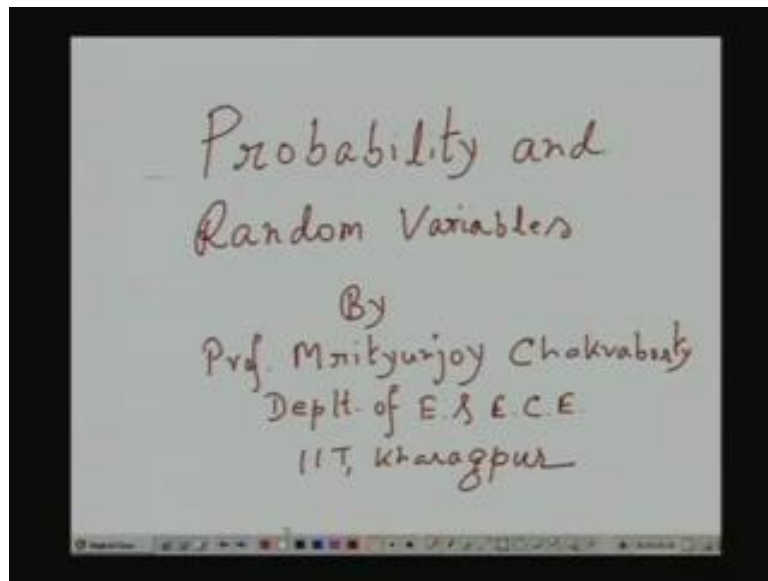


Probability and Random Variables
Prof. M. Chakraborty
Department of Electronic & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 1
Introduction to the Theory of Probability and Random Process

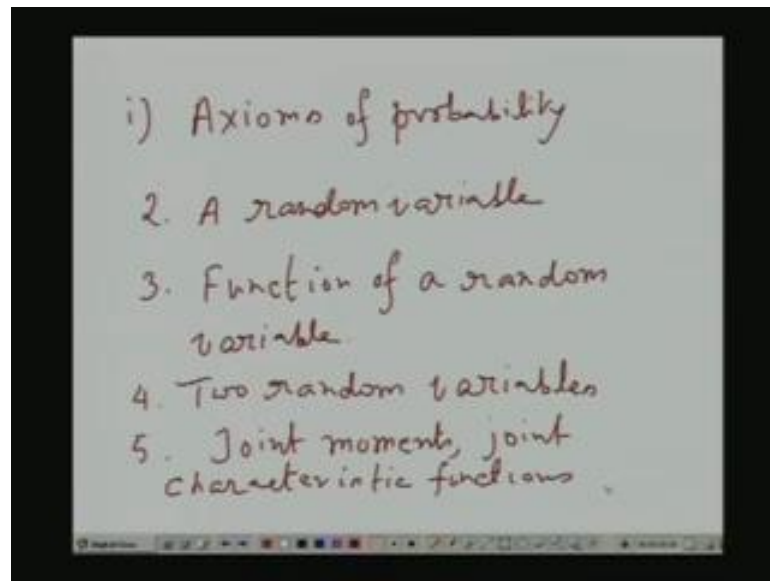
So, I first welcome the students to this course on probability and random variables. My name is Mrithyunjoy Chakraborty. I am a faculty member at the Department of Electronics and Electrical Communication Engineering at IIT Kharagpur. Well, purpose of today's lecture will be very general. I will first give you an overview of the top of the course and that itself will be somewhat vast, and having done that I will devote the remaining time to introduce the students to the concept of probability from an exhibiting point of view.

(Refer Slide Time: 01:40)



As regards to the overview of this course, you know, we will first start with probability. We will first start, I have to start with probability. We will have this topic.

(Refer Slide Time: 01:53)



Actually, I mean, probability is often defined in common practice, you know, I mean from a frequency interpretation point of view. That is you just have so many trials and you observe a particular thing so many times, and just based on that you talk of some kind of estimate of a chance of occurrence and all that, but the theory of probability requires much more than that. You know it needs what you call axiomatic definition and that definition relies on set theory concepts. So, I will deal with that and there by develop the idea of probability. And then that will actual involve some important concept called probability space. After having developed the motion of probability, I will take you to another very important topic that is called conditional probability.

After having done that, I will cover random variable; a random variable; a means one; one random variable. So, I will introduce the concept of the random variable there will be general distinct description. And motivation will be to take into these important notions like: probability density function of a random variable and probability distribution and then conditional density conditional distribution. In that connection there is something else called Bayes theory and this Bayes theorem can be used to find out the total probability. So, that will be covered in this topic of random variable.

Next topic will be function of a random variable. Often you do not only deal with the particular random variable in question, but sometimes you evaluate some functions; may be real valued function, may be complex valued function, a function of that random variable. So, even if you talk about the probability density, probability distribution and

things like that for the random variable, how to have the similar concepts for not the random variable, but its function? Like what will be the probability density function for function of the random variable? The function will be given. Once you do that, that will help to you talk of some very topics called moments.

Moment means, I mean, when you raise the random variable, may be the random variable say the real value random variable, to make life simple, you raise to some power say power p ; p is an integer and take its average; that we will call p th order moment. So, these moments are very important. At least you can understand that for first order moment what you get is very useful quantity; this is called mean. And for second order moment, its leads to something called variance and then correlation and things like that. So, all this can be covered by using the concept that come under this function of a random variable.

Then, there is something very important called characteristic function. Given the random variable of some particular probability density function, there is some function actually it is related to the Fourier transform or the probability density and that is called characteristic function. That characteristic function can be used to find out probability density of a set of random variables; may be say random variables, when n number of deferent random variables are summed, what will be the probability density of the sum? So, that is for this function of one random variable. This will take us to the concept of two random variables.

Actually, whenever you have got one random variable, you know, you describe its statistical behavior by finding out of probability density and things like that. Alternatively, it can be shown that you can describe it completely by specifying all its moments: first order moment, second order moment, third order moment and like that, but that is fine as far as one random variable is concerned. But you know often in practice in real life, we will find that there are many parameters related to particular system or a process or operation which are random, but they have their mutual relation also. They are mutually dependent on each other, mutually related to each other and all.

In that case, you often face a situation where you have to take a call. You have to just face; you have to take into account the entire set of the random parameters or random variables and not just one because behavior of one is not governed by itself alone, but governed by other random variables which are related to it. So, naturally, in this case,

what you talk of is not single statistics, but joined statistics; joined means collecting all of them together.

Now, it is a first step towards that. We will now consider that. We will then consider case of two random variables; that is suppose there are two random variable x and y . So, so far we discussed, I mean, earlier the probability density and distribution of a particular random variable, single random variable. Same concepts will now be extended to the case of joint random variable.

Now, of course, I forgot to mention that I mean, when I discussed this concept of random variable and all that, and especially you know I mean I discussed this probability density and distribution and all that, I will be basically also presenting something useful; density functions which come which we come across frequently in life, but even if we do not come across in life, they turned out to be very useful in modeling phenomenon. One of them is Gaussian density; then there is something rally density; there is something called Poisson density and uniform of course, it is very simple and things like that.

When I come to these two random variables, I will again extend to those univariate. We call it univariate densities. So, univariate Gaussian density to a bivariate; bivariate means when you have got two random variables, bivariate density. So, those examples can be extended to a two variable case. And again, we will have to also define function of two random variables. Like earlier, I have taken here function of one single random variable. It can so happen that your system has two random variables mutually related and you are evaluating some function based on both.

For instance, suppose there is a school and there are students. You are finding out age of the student. That is random because that varies from person to person. And may be the IQ level of student; that varies from person to person, but they are somewhat related. Mentally a five year old kid will not have the same IQ as a twenty year old student, but based on these two parameters, evaluate some function and that function has some numerical value. So, I will call it a function of two random variables.

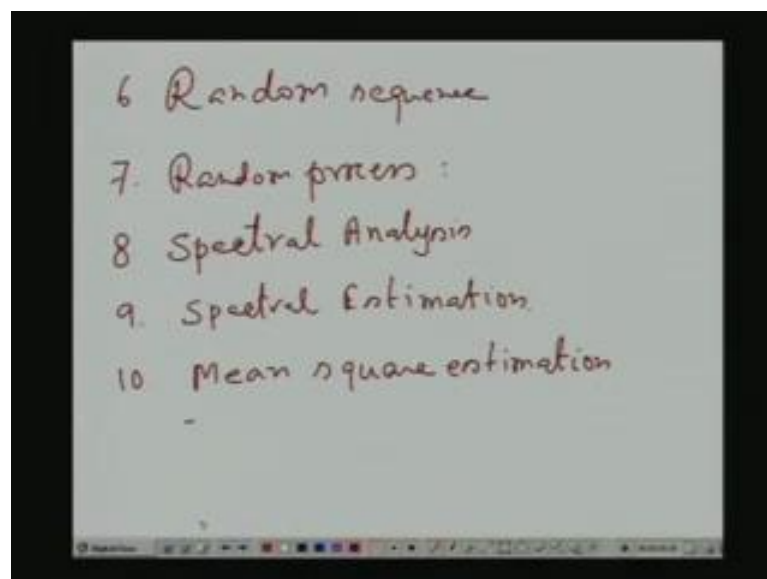
So, again, whatever we have done for function of one random variable, we have do that for function of two random variables; that is finding of the joint finding of the density; probably the density function for this basically. And after doing this, there are certain topics not particular for a particular chapter of the book, but certain selected topics from various chapters which I feel are important and will be covered. That is, earlier we

discussed moments, first order, second order and so on, for a single variable. We have to do it for the joint case also, that is joint moment. And similarly joint characteristic function and this will also be useful to introduce you to the concept of mean square estimation which I will deal with later.

There is something called Hilbert space or other vector space treatment of random variables which is basically geometric, which gives you a geometric interpretation of the whole process of estimation; that will be developed there. I will also touch upon another topic called random sequence; that is a sequence of random variables coming. I will just touch up because that itself is a very extensive subject you know. And in that context, what is important is this notion of stochastic convergence. That is, we know that is, any sequence can be converging or diverging; if it converging, it converges to a limit under some condition. There is some called Cauchy condition and there are many things you have studied.

There are converging tests also, but when it comes to random sequence, convergence motions have to be defined because if the particular sequence is not important, the class of sequences, they should converge to a point. So, that relates us to stochastic convergences; that is very important. So, I will touch up on these topics after this. So, I just mentioned them here, because I am writing I consider this joint moments, joint characteristic functions. And the other topic is other topic that was number 5.

(Refer Slide Time: 12:27)



The other topic is number six is Random sequence that is which I just mentioned last. So far, whatever you have covered you know up till now. I mean under this plan of course, we have not covered anything; that is a different matter, but under this plan whatever I have we have touched upon. We basically relate to first probability notion, then probability density and all that, and then takes you to individual random variables. Mostly we considering continuous valued random variables, but the concept can be easily extended to discrete valued random variable also. They are called discrete random variables. That also will be done, but after having done these, we have to extend this concept of randomness, this random variables to something called random process.

Random process means we will be now considering waveforms which are random, not just a variables, waveforms, and waveforms could be a continuous wave form or could be discrete sequence. That relates to the concept of random process because this random process is which you finally come across in life.

Any observation that you take, especially in engineering, you know that is only some period of time. You observed you have a wave form and that wave form can vary from experiment to experiment; from observation to observation. So, that is what we have here in the form of random process. So, I will be dealing with this topic after this; random process. Here, we will first have first introduce you to the general concepts of random process and then there is something called stationary process. Stationary process means randomness. It is homogeneous, distributed along the axis, you know I mean of any bias towards a particular location of the time axis. I mean, we will of course discuss at length when you are going to those topics that is called stationary process.

Again stationary process is a concept which is often not make in practice, but we will take this how to do to that, because otherwise analyzing things or you know devising algorithms, or processing means becomes very difficult. And again, another concept is related to stationary; that is called Ergodicity.

Ergodicity comes from the need rather, I would say because we often assume the process to be ergodic because the need is such that I mean you may not get several versions of the waveforms in real life. In real time, you are just getting one waveform and based on that only you have to make estimation. Using the motion of ergodicity, it has three device, your processing algorithms and all that to process that actually. I mean you can then there by bypass the need of having several versions of the version; I mean the

several observations of the same waveform to make say average or things like that; that requirement can be bypassed.

And then comes is very important topic that if you pass one such random process, say stationary process, through a system which is linear and time invariant, both discrete time will be considered and continuous time also. Then the output is also stationary. If, so what is the out I mean we will verify that. Of course, stationary, with this stationarity certain motions of correlation and things I mean related things are associated. So, those have to be examined both for the input and output. And based on that only, we can come to those infinite sets. So, basically, that is, that will be covered under the heading Systems; that is Linear time variant systems with stochastic inputs.

After these comes this topic Spectral Analysis. In spectral analysis, we will introduce you to the notion of power spectral density of a random process. So, that we know that needs this notion of stationarity first and also I mean why power spectral density, all that question will be dealt with. Properties of power spectral density will be explained in detail.

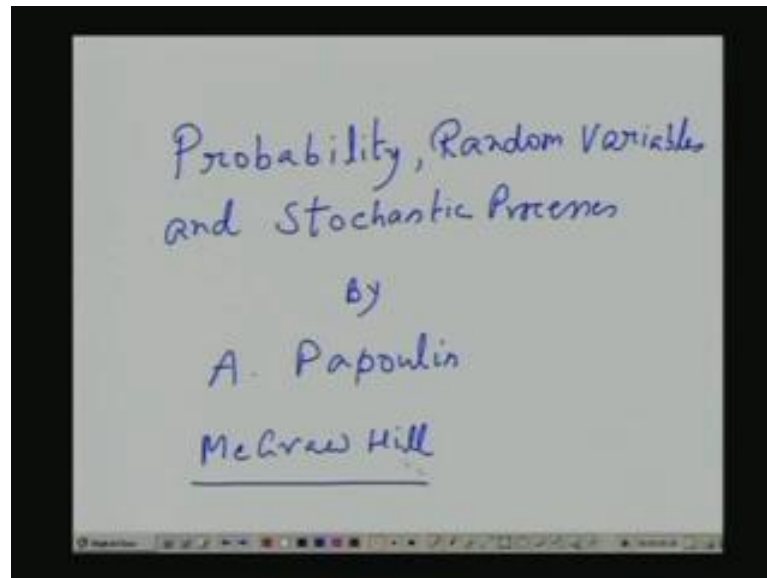
And then, as I told you, that previously I will be discussing the effect of a linear time variant system or random input, that concept will be used to show how you can generate random processes of some particular wave shape, I mean particular power spectral a particular shape; I mean random processes with some particular type of power spectral density, particular shape by designing linear time invariant system.

So, that will be considered. I think up till then, I will be considering fairly I mean good amount of material, but still two topics I feel I should touch up on as much as possible, what I mean, if time permits. One is called estimation of the spectrum that is spectral estimation. Of course this is subject to the availability of time and all that, how far you go. And other one is Mean square estimation. Under this you will be covering topics like linear prediction, smoothing and filtering. Now, I tell you now, this topic spectral analysis, that the spectral estimation, spectral estimation and also mean square estimation, they are very huge topics. You can have one full semester course just on Mean square estimation.

Similarly, for Spectral estimation, there are I mean books and I mean entirely devoted to these topics, but obviously, there is no scope of getting too much into all these, but I will try to give some flavor of these topics towards the end of this course. That is my

tentative plan, but if I find that, you know I mean, as I go along if I find that in the context of whatever I have mentioned, I need to inject more input; I know I you know I need to discuss some more things, then I will introduce them I will add on as and when needed.

(Refer Slide Time: 19:47)



Book though there are many books available, there is one book which I maintain. It is given in here. I write its name. It is called... This is a book published by McGraw Hill. I am sure there are economy additions available. This is a very famous book. A book celebrated all over the world and this particular author is very well known for this particular book, and of course, he has another book which is also equally famous; almost equally famous. It is called Signal analysis, but that is not being covered here; that is not being considered here. So, this is the book, but I again tell you that there are many equivalent books available in the market.

So, as long as these topics are, same topics are, covered, any book for matter will be welcome. So, I know, after have been discussed this, let us now talk in general about this idea of probability. Actually, this topic of probability comes in the context of studying some natural phenomena or maybe you know some experience that you carry out, for we find that, I mean the outcome of the experiment is not something fixed, but there is something random about it. That if you try to observe, the same thing again and again and again, you find that you are not observing the same thing.

For instance, as I told you suppose you go to a school. A particular class, a particular class you go to and then you just ask for the ages, exact ages; year, month, date, number of days and like that for each student. You will find that no two figures are matching. For each student his or her age is something and for other student again you know I mean the same the ages something else. Even though they are close because after all in a class you expect the students to be having age close to each other, but not exactly same. So, here the outcome which is the age of the student is varying.

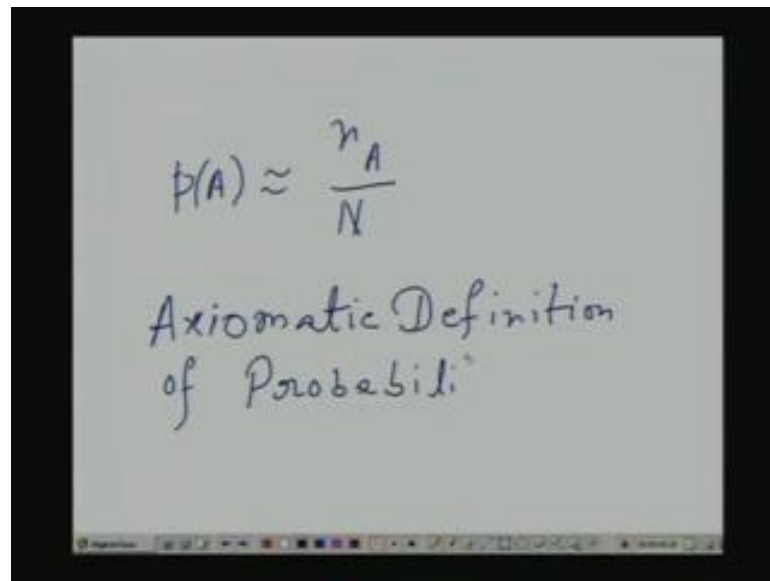
Similarly, suppose there is a chemical process; some process is going on. There is a furnace and you want to measure the temperature of the furnace or some part of it every now and then. You find even if you try to measure, you do not find exactly the same temperature. I mean you find temperature in a certain range, range could be narrow, but the temperature fluctuates. In such contexts and there are many other contexts now you yourself can visualize based on these examples.

In such contexts, one nice thing has been observed that as far as the outcome is concerned, I agree that there is randomness. There is some unpredictability. There is randomness. You call it chaos; you can call it anarchy, whatever, but there is certain thing; there are certain things which do not vary with time. That is, it has been found experimentally that for each of these experiments, certain averages if taken over a sufficiently large number of outcomes or observations, this somehow approach some steady value. And even if your number of observation goes up further and further, this average do not change. So, this average is rather characterized the particular process particular experiment that is going on.

Further, even if you do not take all the outcomes, suppose outcome one, outcome two, outcome three, outcome four, and like that there are sequence of outcomes. Even if you do not take all the outcomes, but take some subsequence of the outcomes, you will still find that the average, that certain average that I mean, that average value remains same. To give an example, suppose you are tossing a coin and I mean as you go on tossing the coin, you find out after a while the percentage of, you know, the percentage of occurrence of head; so many trails and so many times I got head. So, what is the percentage? You will find that if you toss it several times, you know. Then that percentage figure will approach the figure 0.5.

If you go on still tossing, I mean you see that the average will not vary much around 0.5; it will approach the steady value if your number of observation taken already is huge. Same thing will happen even if you do not take all the tosses. Suppose I take only the every 4th toss and see its outcome and based on those outcomes only I do the average. Even then I will find the percentage of occurrence here is 0.5; I mean corresponding percentage. So, that leads us to this motivation for concept of probability.

(Refer Time Slide: 26:09)



The image shows a whiteboard with handwritten text and a formula. The formula is $P(A) \approx \frac{n_A}{N}$. Below the formula, the text reads "Axiomatic Definition of Probability".

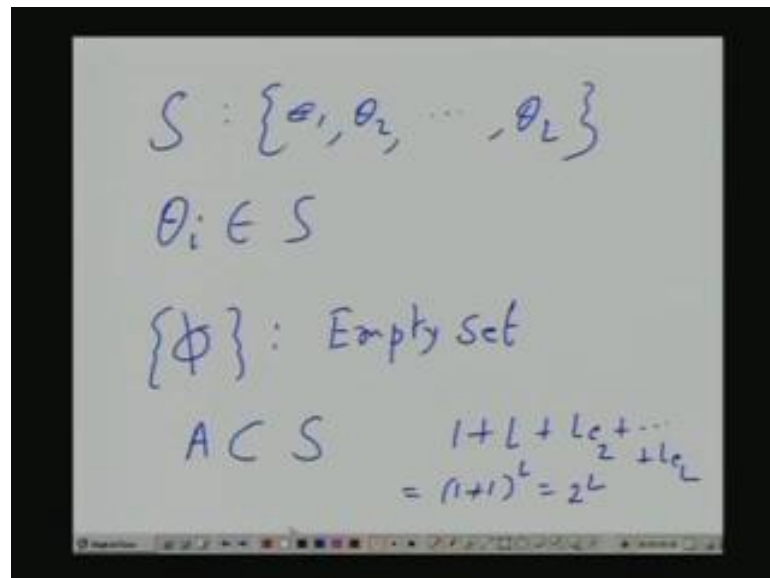
In the case of probability, classically, if there is any experiment that is going on and you are observing outcomes and total number of observation is capital N , out of which a particular event occurs, particular event say A , that occurs small, I mean, these times you know n subscript A times. Then the corresponding probability is nothing, but this. This is a classical definition. It is called frequency interpretation of probability. It gives you a chance of occurrence of A , provided of course N is large, number of observation is large. I have taken so many samples and out of which so many out comes out of which only on n_A occasions I got A . So, if capital N is large, then this ratios does not change its value even if you take more and more out comes. So, that is a classical interpretation of a probability. It is called frequency interpretation.

So, in practice often it comes to help, but in the theory of probability, you know, we will just we do not rely on these. We try to develop a mathematical definition of probability. When you have to really estimate from real life data, that time you can take recourse to this, but we want to evaluate its properties so that we can device algorithms. We can

device processing means based on these properties. That take us to what is called the Axiomatic definition of the probability.

Now, to get into this, first thing we have to do is to revise the set theory concepts because this axiomatic definition is based on set theory motions. So, I will quickly go through this and then maybe I will just take to the stage of, we stay where we can formally get into this executing definition, but that detail discussion on this executing definition of probability we will carry on in the next class.

(Refer Slide Time: 28:39)



In set theory, first we have got say set S. And the set S, to start with, it consists of some elements. Elements are say theta 1, theta 2, dot, dot, dot, dot, say theta L. This is a collection of elements. This is a discrete set. Discrete set means it has it is a countable discrete set. Countable because you can count that total L elements and discrete because either this or this or this or this.

I start with this sort of set because this is very easy to understand, but this set could be a continuous set. For instance, you know, this set could be segment of the real line between the number 2 and 3; Is it not? And that is not countable; you cannot count the total number of points that you will have between 2 and 3. Sometimes a set can be countable, but still infinite. This is a finite countable set. This is very easy to understand.

Here, here, or for any set as you know I can write theta I element of S. This stands for element. So, it is an element of S; that is this is member of the set S. Then there is something called, this will be called empty set. And subset, any subset of S is a

collection of the members of the set S. So, each member is a subset; this empty set is a subset.

Empty set is a subset because empty set is already a part of it; empty set is you know part of any set. So, this is a subset of this. A set consisted of only one element theta 1 is subset of S. A set consists of any element 2 is a again subset of S; likewise or may be a set consisting of theta 1 and theta 2 is a subset; theta 2 and theta 3 again is a subset. And likewise, if A is a subset of S we will write like this A is a subset of S.

Now, can you tell me for a set like this, how many subsets in total can you have? 2 to the power L; that is the correct answer. Why?

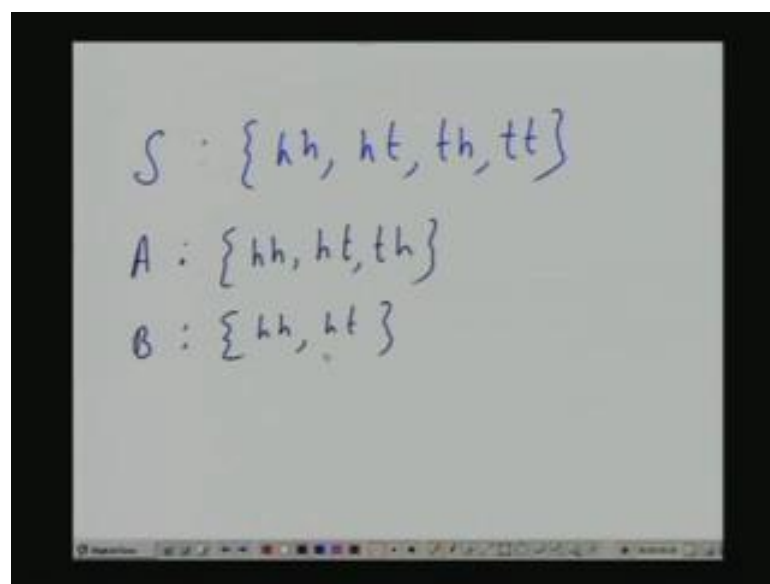
Please tell.

Student: ((Refer Slide Time: 31:29))

Actually it is like this. I mean you start with 1 that is this; 1 that is for the empty set; then 1, 1, 1, so L subsets. Then you take 2. So, that means, how many combinations? $L C 2$.

Then $L C 3$ and like that up dot dot dot dot upto that $L C L$; that will be nothing, but a binomial expansion of 1 plus 1 whole to the power L, sorry whole to the power L which is 2 to the power L. So, I have got so many those many subsets. For example I will just take an example from the book because that it is very good in the context of our business.

(Refer Slide Time: 33:00)

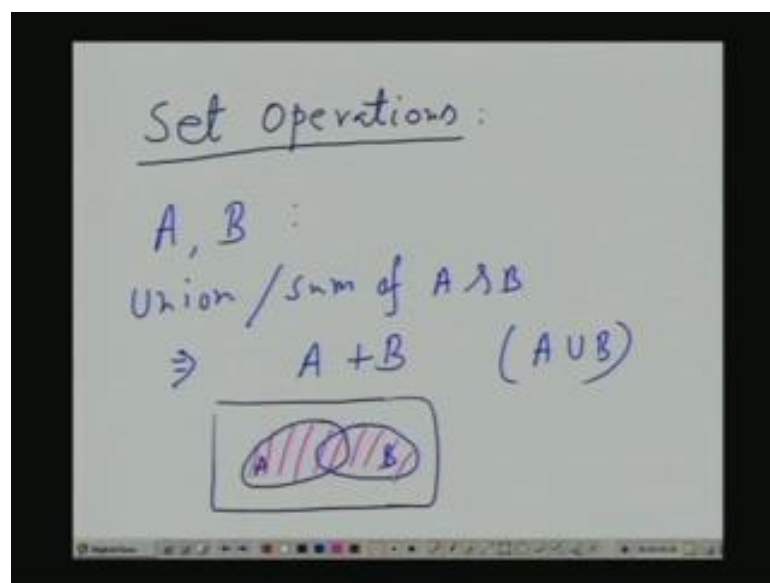


A photograph of a whiteboard with handwritten mathematical sets. The sets are defined as follows:

$$S : \{hh, ht, th, tt\}$$
$$A : \{hh, ht, th\}$$
$$B : \{hh, ht\}$$

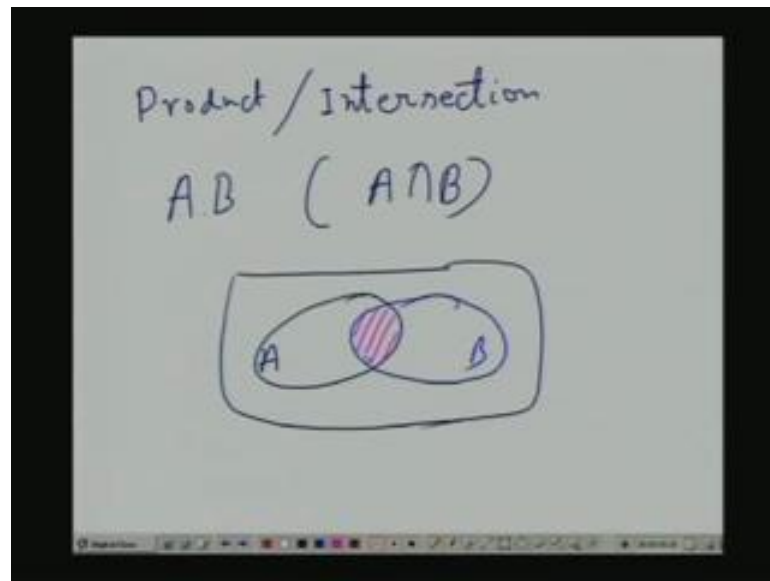
Suppose I toss coin. Coin is tossed twice and based on those two tosses I have a set S which I will write like hh, ht, th and tt. Here, hh stands for that event where both on first and second toss you got heads. Here, first you got head; then tail; tail and then head, and both tails. In that case, I can have a subset say A as ((Refer Time: 33:42)). What is a subset? It is a subset of those events, these are all events, is subset of those events where at least one head as occurred. Similarly, I can have another subset. These are subset where first toss always give you h and likewise. That takes us to set operations that I got to do in the next page.

(Refer Slide Time: 34:54)



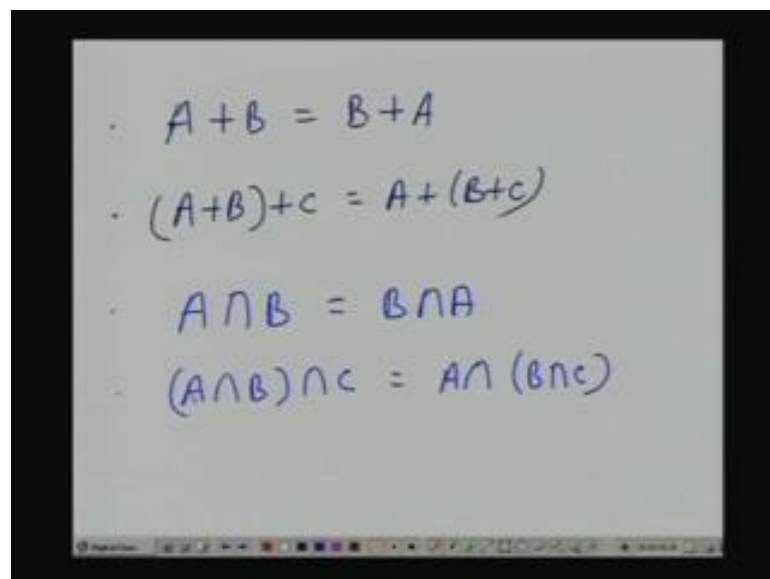
One is union giving a set A and B , two sets A and B , I define union or sometimes we call it sum of A and B as A plus B . You are used to this notation A union B , but as I am I will be going along with this book, I will follow this plus rather than this; this not arithmetic plus. What does it mean? We all know, but still it is worth repeating that resulting one is a set where all elements of A and all elements B are present. So, if you have a Venn diagram, you know, if you have a Venn diagram like this, suppose this is A , this is A and say this is B , then A union B will be this entire toss. This is A union B , we call it A sum B also.

(Refer Slide Time: 36:51)



Similarly, product or intersection, this is denoted by \cap . You can put a dot, but you are possibly used to this notation, but again going by the book I will follow this. This means this intersection is a set where only those elements which are present both in A and B are content. So, if basically looks at the overlap between A and B. That is in terms of the Venn diagram, if you have got this to be A and this to be B, then this is your A union this part is sorry A intersection B, this part.

(Refer Slide Time: 38:45)



Now, to the next one, this union and intersection; they satisfy some properties which can be easily seen; that is you know coming to union first is same, as this is obvious. Union

of A and B or B or A. This is called commutativity. Then other thing is again no need to draw a Venn diagram for this. This is I mean this can be easily in fact, from common sense. That is you are taking a bigger set here where all elements of A and B are present. You are doing the overlap; that again and again you are summing with C. So, resulting thing will contain all the elements of A, B, C; same will be here.

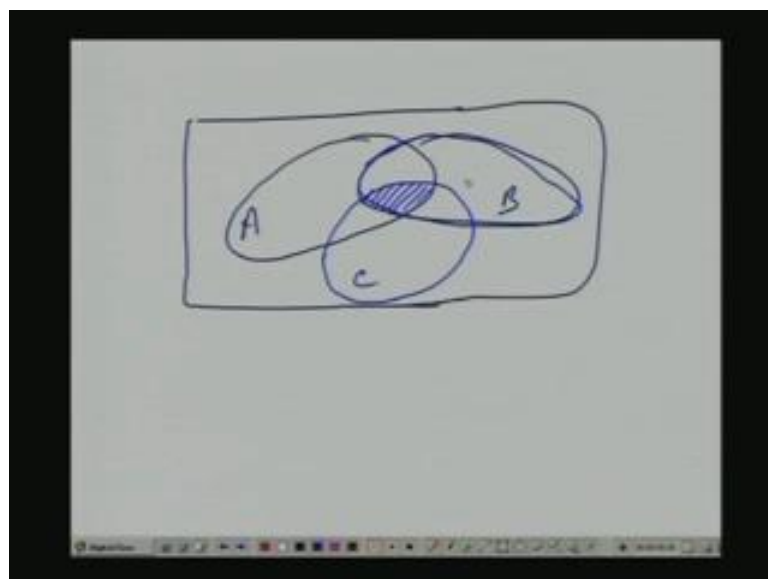
This contains all elements of B and C, the bigger set. With that if you combine A, then again total set will be what? It will consist of all elements of A, B and C. So, that is a both sides are same. Now, when it comes to this, when it comes to product, this relation is very simple. I mean you can easily understand what it is. Overlap between A and B or between B and A; they are same. It is just to define sentences.

How about this? Is it same as ((Refer Time: 40:45)). In fact, I am using this intersection, you know, but actually interchangeable. I will use it in the productions also that is AB; AB is BA and AB C is A B C.

Student: ((Refer Time: 41:18))

No. No. This is associate like A B with C is same as A product with BC. So, here A plus B plus C. These two things are same. This is called associativity. There we can see this just by this Venn diagram.

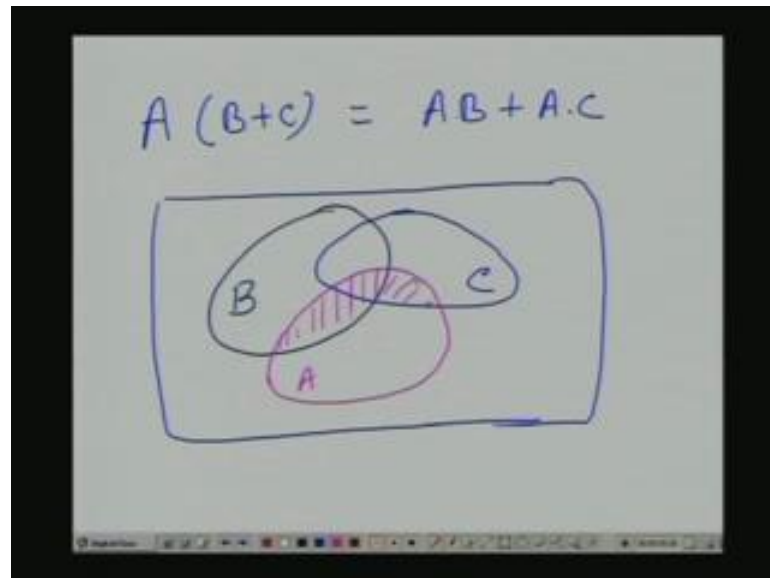
(Refer Slide Time: 41:40)



Take again thing like this: A, this is C; in fact, let me use this. What is the intersection between A and B? That is this entire region? That intersection with C will be what? This

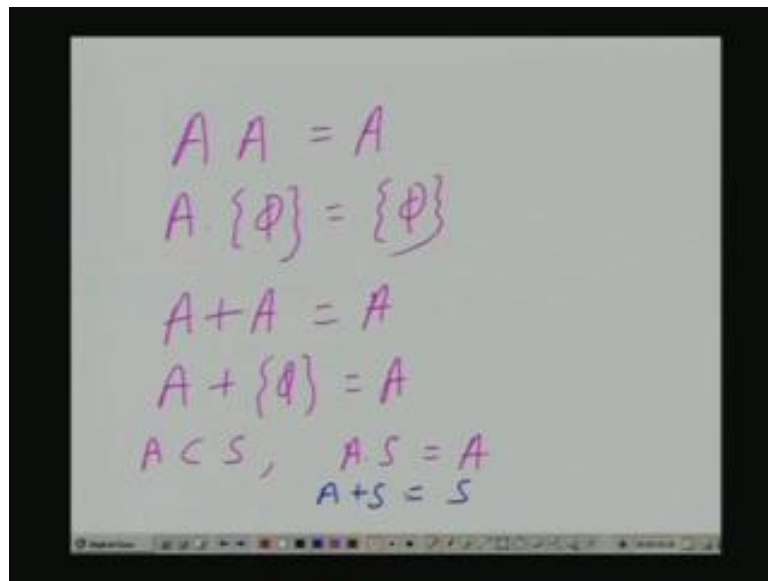
part, so what is A intersection B or A product B, and they product with C? That is actually the overlap between A B and C; the common part that is this part. Similarly, if you consider B product C first, that will be this entire region. That intersection with A will be what? Only this; this is to be the same.

(Refer Slide Time: 42:57)



Then comes other property which is a combination of both linear and intersection; that is I am using this style here instead of using the intersection symbol; the product symbol is here. Using the same thing this will be... Again, for that, we take this. This is A. Let me erase it; let me call it B; this is C. So, B union C is this entire region that intersection with A. Now, suppose your A is this. This is your A. So, what is A intersection B union with C? B union with C is this entire region; with that A is intersecting. So, that means, this zone. What is A product with B, that is A intersection that is this part? What is A intersection with C, this part? And their union will give you again this. So, this is actually called distributivity that is this intersection is distributing or union.

(Refer Slide Time: 45:48)

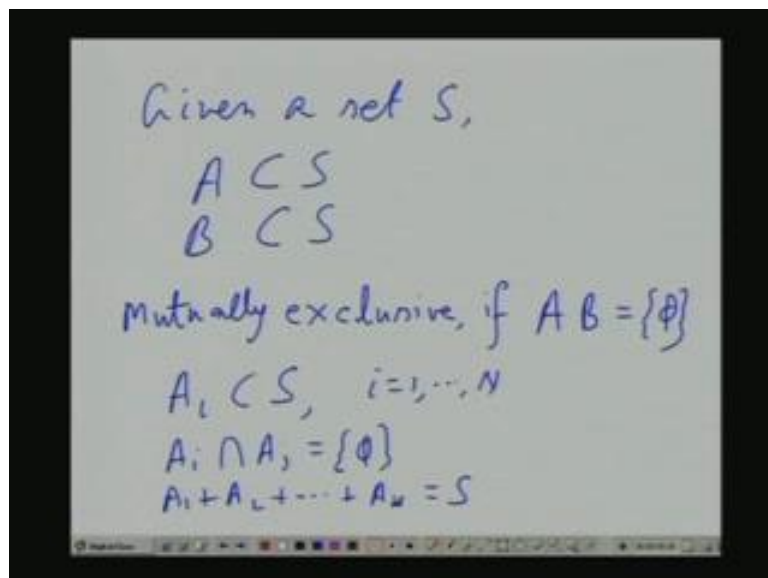


A whiteboard with handwritten mathematical identities in pink and blue ink. The identities are:

$$A \cap A = A$$
$$A \cap \{\emptyset\} = \{\emptyset\}$$
$$A \cup A = A$$
$$A \cup \{\emptyset\} = A$$
$$A \subset S, \quad A \cap S = A$$
$$A \cup S = S$$

Certain things you can see that A intersection with A will give you A. A intersection with phi will give you what? Phi, because this is common in this also A plus A, A union with A is A; A plus phi, here is a difference, that will be A. In general if A is the subset of S, then A intersection with S will be A itself and A plus S will be S. It is all elementary things.

(Refer Slide Time: 47:05)



A whiteboard with handwritten definitions in blue ink. The text reads:

Given a set S,

$$A \subset S$$
$$B \subset S$$

Mutually exclusive, if $A \cap B = \{\emptyset\}$

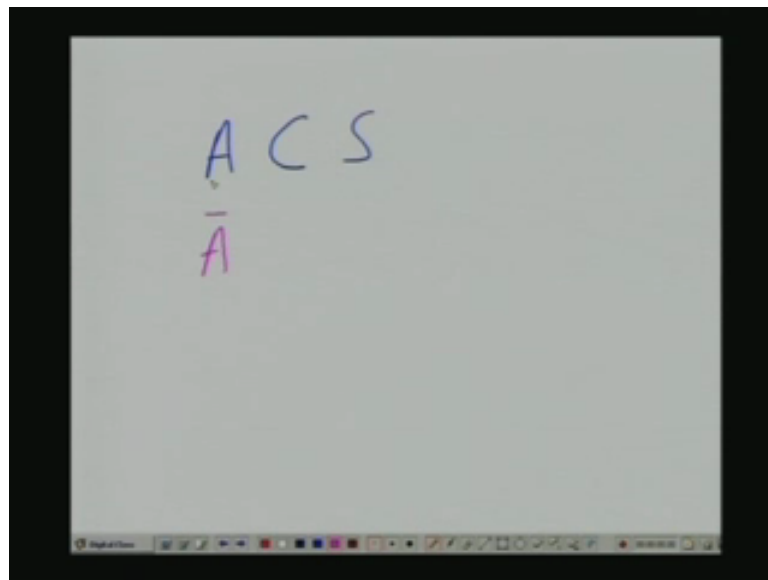
$$A_i \subset S, \quad i=1, \dots, N$$
$$A_i \cap A_j = \{\emptyset\}$$
$$A_1 \cup A_2 \cup \dots \cup A_n = S$$

Few more definitions. Suppose you are given a set S. You are given a set S. They have got two subsets A, B. A and B will be called mutually exclusive if there is no intersection.

Using probabilistic motions it means that A and B could be two events which cannot occur simultaneously.

Just took an example. This takes us to concept of partition. Suppose there is a set S. Out of the set S, we derive some subset A_1, A_2 , up to say A_N such that A_i , each A_i is element of A_i subset of S; i equal to say 1, dot dot dot dot N and A_i another thing is they are mutually exclusive. $A_i \cap A_j = \phi$. Another property is they are taken such that if you take the union of them, you get S; that is another thing; $A_1 \cup A_2 \cup \dots \cup A_N$, that is your S. Then this set A_i is called the partition; that is intersect S has been partition into smaller subsets so that no space has been left out. You just invite them by union, you get the intersect S back, and there is no intersection between them.

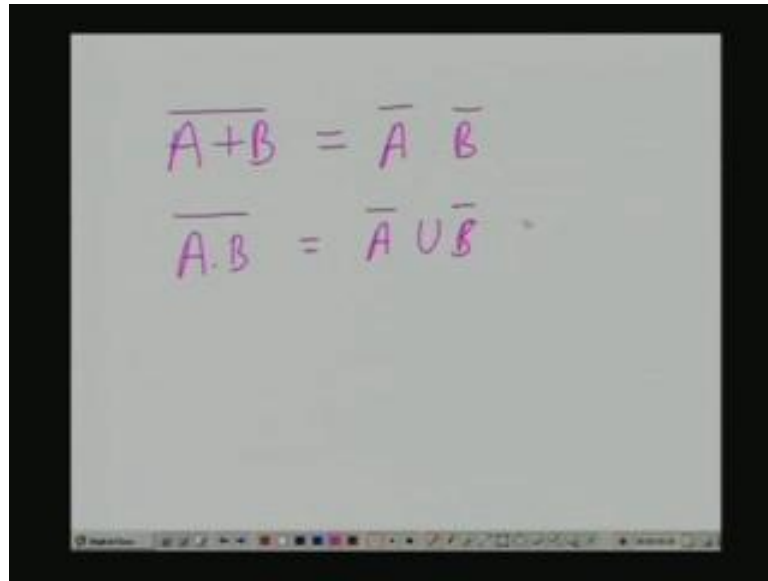
(Refer Slide Time: 49:28)



And last thing is compliment of a set. If A is subset of this, then A bar is called a compliment. A bar means all elements of A which are not present in A. That is called compliment. So, if you take capital S itself, then its compliment is empty set and empty has its compliment equal to the total set S.

So, this takes us to... Well, before you are going into that, maybe we can just quickly tell this Demorgan's law which you have already studied, you know, in the context of digital circuits where again they talks in terms of logic, but it again, you know, underlying principle remains same. Demorgan's law is like this, you know, I mean if you take this, you can verify by Venn diagram. I give it as an exercise.

(Refer Slide Time: 50:32)



The image shows a whiteboard with two mathematical equations written in purple marker. The first equation is $\overline{A+B} = \overline{A} \overline{B}$ and the second equation is $\overline{A \cdot B} = \overline{A} \cup \overline{B}$. The whiteboard has a black border and a small toolbar at the bottom.

I will take compliment; that is take the union of A and B; its compliment. You can verify by Venn diagram. That will be nothing, but intersection between A compliment and B compliment. And on the other hand, if you take the intersection and then take compliment, this will be nothing, but... sometimes I am using this union, sometimes plus. So, please do not get confused because I am more used to this union symbol and I am sure you also.

These are the two main Demorgan's law. I am not getting into the proof. You can easily do it using set theory concept using the Venn diagram. So, this is as much as, as far as this set theory part is concerned. This will take us to the concept of probability space.

I will discuss this at length in the next class, but just to give you an idea about why we deal with so much of set theory today is that you know I will first define some set which is called the probability which actually called as space. That is space will be the total set and that will be called the certain event; that is that that will always occur. Some experiment is going on. So, that set will resemble, will represent something which is bound to occur. In such an event, there is a total set. Each member of that set will be called an experimental outcome; that means, the set will consist of all possible outcomes.

Therefore, the set is an event which is bound to occur. Those outcomes, you can collect some outcomes and form a subset. Each subset will be called an event. So, event can be each of the outcomes or a combination of the outcomes or the entire set itself. And then on that set will be, you know, given a set S like that and subsets, there we will introduce

some measure, some function or measure called probability. That will be the axiomatic treatment of probability.

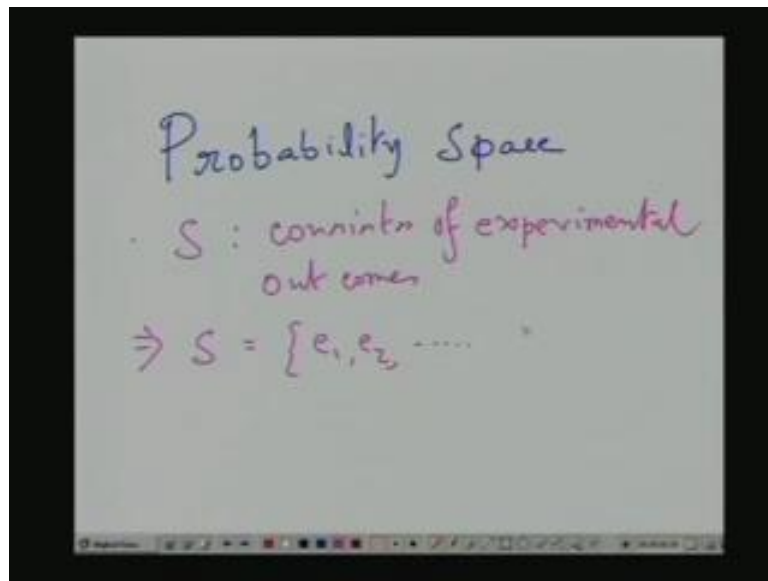
So, I will be stopping here today. Just to summarize what you have done today is this. We have given a first overview of the course, but even though I try to be I mean exhaustive to put the point there, I added that you know I mean as I go along, I might introduce certain extra things in each chapter which I have not named today. And after having done that, we have discussed the idea of probability is a need or motivation in general terms. And then we have explained that this is a classical frequency interpretation based notion of probability alone is not enough because to take a proper I mean we need to give a proper treatment to this topic. And that takes us to the axiomatic definition for which set theoretic knowledge is required because using some sets only we will develop these axioms. That we will begin from the next class. So, if you have any questions, fine; otherwise, I will wind it up. That is all for today.

Preview of next lecture:

So, in the previous class, we had first discussed we have presented this outline of this course. Detailed outline was given as to what all will be covered and then I tried to develop the motivation for this theory of probability in very general terms. And at that time, I mentioned that this theory will be developed using some solid beneficial foundation of set theory, classical set theory. And then I discussed the basic set theory notions, definitions and operations at length.

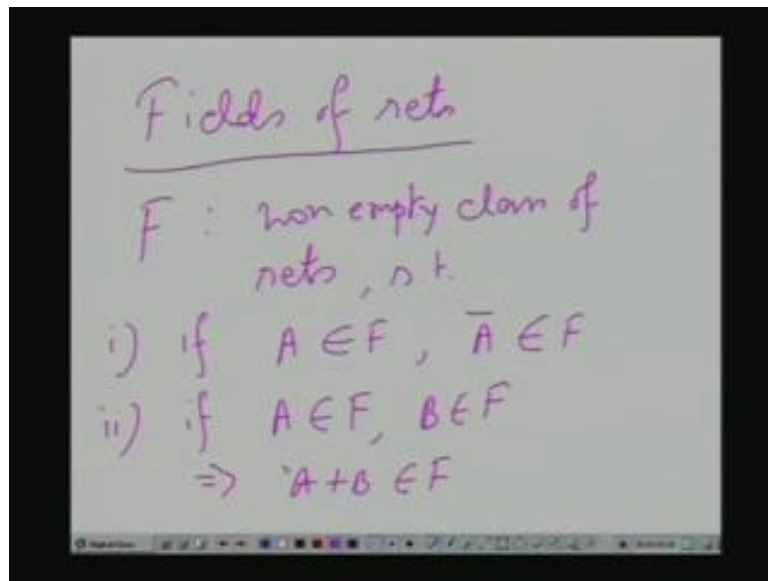
Today, I will be using those concepts of set theory to first introduce you to the topic of the subject of probability space and we will be discussing the properties of probability space. And using this definition and properties, I will then go into what is called axiomatic definition of probability. We will call it axioms. And then, we will get into the consequence of those axioms and then I will try to relate them to real life world because after all the axioms are purely mathematical. And then, once those axioms are given or described, then I will get into some further sophistication in the form of something called field of sets; in fact, a particular class of fields, particular type of field called borel field. And in this borel fields only, mathematically we can apply this set theory with this probabilistic axioms correctly. So, I will go up to that. We will also consider some examples.

(Refer Slide Time: 55:58)



So, first I start with probability space. First, let there will be a set S . This set can be finite, countable or infinite, finite and countable at the most you know the simplest of all cases. Finite there is total number of entries is finite and countable. Then it can be infinite, but countable like, you know I mean, do you know we can you give an example of such set where the set is infinite, but countable? Say set of integers is countable is called countable infinite or set can be itself, non-countable infinite where is real life. It can be anything. But for our treatment, I will try to first take those examples where S is finite and countable. And then we will try to get in fact how do you try to generalize it to those infinite sets. Then only I have to get into that what I said field of sets and borel field and all that.

(Refer Slide Time: 57:44)



This is set S . It consists of experimental outcomes, I will explain, meaning S is something like e_1, e_2, \dots . Now, I am constructing a field. If A is present in that field, A bar also must be present. And number two is if, this should be actually element, because these are now sets are now elements of the class. Class is what? A collection of sets. So, each set is an element of the class; it should be element, not subset.

If then, these two, using this we will show that this applies true for this also, the intersection. Using these two, we will develop some property. We will show that if A is an element of F and B is an element of F , then $A \cap B$ is also an element of F .

So, in the next class, I will take from here. And then I get into something called borel field and there I give the more exact definitions of actually borel field. I will tell you why borel fields are required because when you have got that infinite set, all subsets do not qualify for events. Only those which give rise to borel fields, they give rise to, they qualify for events. After having done that, I will get into again more practical things that is conditional probability and all that. We will solve some problems also. So, that is all for today.

Thank you.