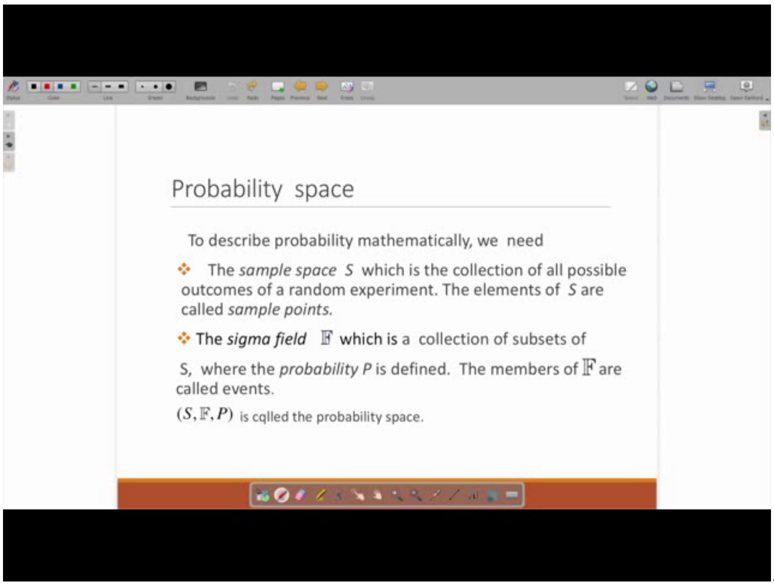
Advanced Topics in Probability and Random Process
Prof. P.K. Bora
Indian Institute of Technology Guwahati

Lecture -01
Probability Basics

In this lecture I will cover the basic probability concepts, these concepts are needed to understand the topics covered in this course.

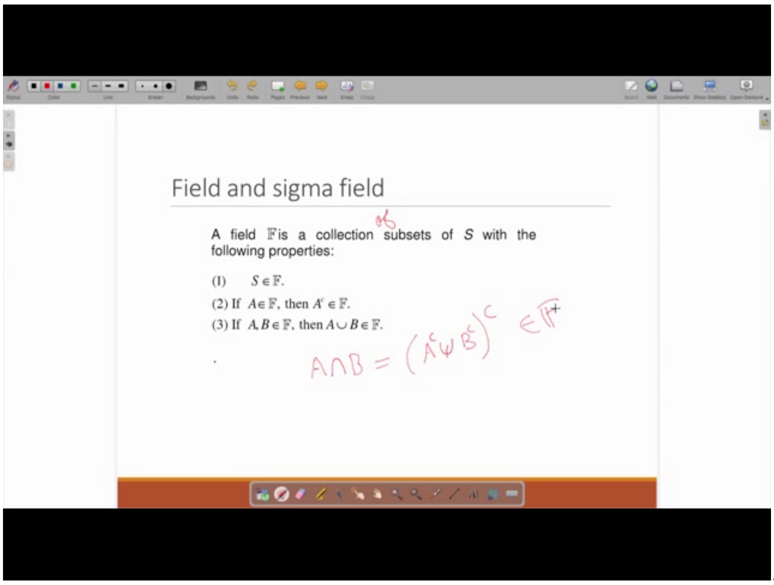
(Refer Slide Time: 00:35)



I will start with probabilities space. You know that to describe probability mathematically we need a sample space S. This is the set of all the possible outcome of a random experiment for example, if I toss a coin that is the random experiment because the outcomes cannot be determined beforehand. So, the possible outcomes are head and tail and the sample space will comprise of the members head and tail.

To define probability meaningfully we need a class of sets known as the sigma field, this sigma field F we see the collection of subsets of S where the probability P is defined, the members of F are called events. We will elaborate these points. Therefore, to define probability first of all we will have the sample space then we will have the sigma field F and probability will be defined on the members of the sigma field the triplet S, F, P is called the probability space.

(Refer Slide Time: 02:07)

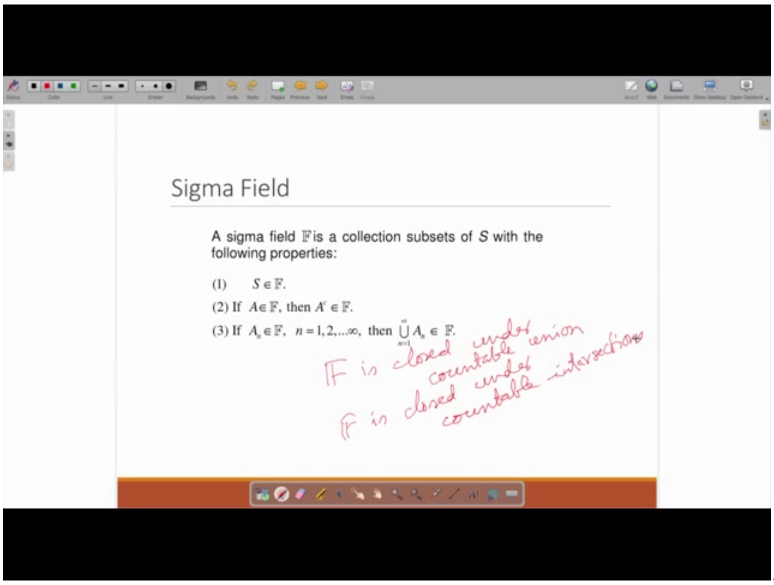


I will give an outline of the concept of sigma field first I will introduce A field. A field F is a collection of subsets of S with the following properties, number 1 this sample space S belongs to F, number 2 if A belongs to F then A compliment belongs to F, number 3 if A,B belongs to F then A union B belongs to F. So, therefore, the sample space is a member of F and if A is a member of F then it is compliment must be A member of F thus F is closed under complementation.

Similarly, F is also closed under union operation, now these 3 properties can be used to derive other properties for example, the field F is closed under intersection operation, how we can show this, we can show that A intersection B is equal to compliment whole compliment of A compliment union B compliment that is the De Morgan's law. So, A intersection B is equal to compliment of A compliment union B compliment.

Now, a compliment belongs to F thereof B compliment belongs to F therefore, there union belongs F and there compliment that is A intersection B belongs to F. So, that way in that way we can derive the other properties set theory properties set theory operations like intersection set difference etcetera can be thus the field F is closed under complementation and union operation and it is closed under set theory operations like set difference etcetera, also an even is a member of the field any member of the field is called an event.

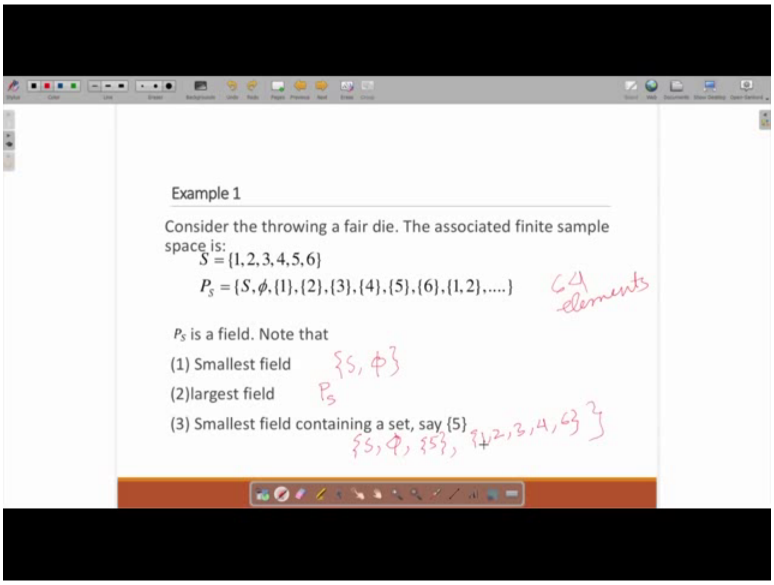
(Refer Slide Time: 04:20)



The field concept is sufficient to define probability in a finite sample space, but when we have an infinite sample space like the real line itself, then we have to introduce the concept of sigma field or sigma algebra field is sometimes called algebra, similarly sigma field is sometimes called sigma algebra. Now first properties of field and sigma field is common only third property is different, we see that if A n belongs to F for n 1 to infinity then union of A n from n 1 to infinity also belongs to F. So, that means, F is closed under this sigma field F is closed under countable union operation. So, F is closed under countable union.

So, sigma field is closed under complementation the entire sample space is a member of this sigma field and it is closed under countable union. So, we can similarly using the De Morgan's law, we can show that the sigma field is closed under countable intersection also, F is closed under countable intersections this sigma field is necessary when we define probability on an infinite sample space.

(Refer Slide Time: 06:30)



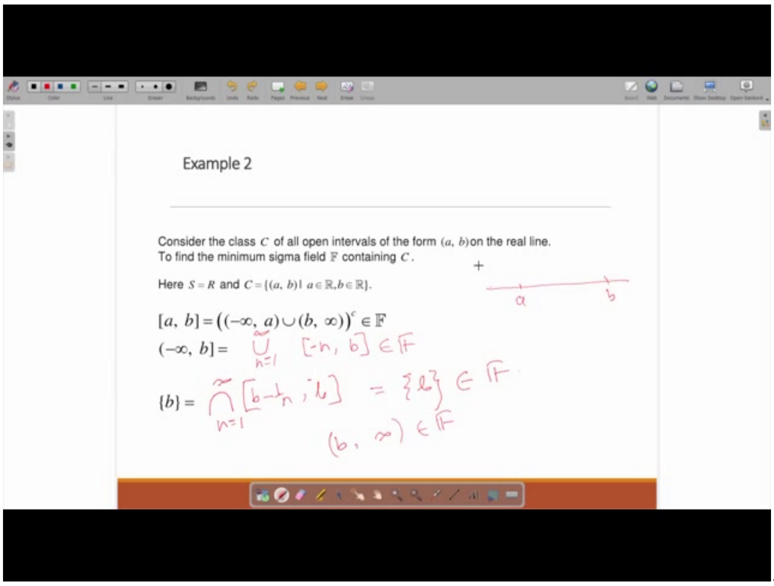
I will start with an example considering the random experiment of throwing a die, the associated it is a finite sample space the associated finite sample space is S is equal to it contains 1, 2, 3, 4, 5, 6 corresponding to 6 faces away die. Now the power set which is the collection of all subsets can be enumerated for example, S will be a member of this power set phi is a member, similarly I can consider elements singleton set like 1, 2, 3 etcetera and double ton set like that. So, in that way how many elements will be there in P S to refer 6, refer 6 is equal to 64 elements. It is easy to verify that P S is a field because if we consider any 2 members their union will be a member of P S similarly if we take the compliment.

For example if I consider 3 it is compliment will be this set consisting of 1, 2, 4, 5 and 6 and it will be a member of P S. We will we also note that this small S field what is this small S field of the subsets of S, if we consider the class like this S and phi clearly S union phi will be S phi union S will be S. So, it is closed under union operation compliment of S is phi compliment of phi is S therefore, it is closed under complementation and is a member of this collection therefore, this is a sigma field or F field because it is a finite sample space you can call it a field and this is the smallest possible field.

Now, what is the largest field because P S contains all the subsets therefore, P S is the largest field, number 3 is smallest field containing a set shape set is given by single element only that is a set consisting of element 5. Now we can enumerate the sigma field or the field containing the set first of all S is a member of this collection, phi is a member of this collection, 5 is a member of this collection and it is compliment that is the set consisting of 1, 2, 3, 4 and 6 that is also member of this collection. Now this if I consider this collection of subsets it satisfy all the properties of the field. Therefore, this is the smallest field containing the particular shape consisting of element 5.

We will consider another example this time we will consider example for infinite sample space.

(Refer Slide Time: 10:07)



.

Example 2, consider the class C of all open intervals of the form a, b this is the open interval on the real line. So, we have consider the open interval like this, this is a, b greater than a, and this interval excluding the end points a and b. These are the this is an open interval similar open intervals we are considering to find the minimum sigma field F containing C. So, we are considering C the collection of all open intervals to b a sigma algebra what are the other sets that we have to include the in this collection. Here S is R and C is defined by this set that is the open interval a b, where a belongs to R, b belongs to R.

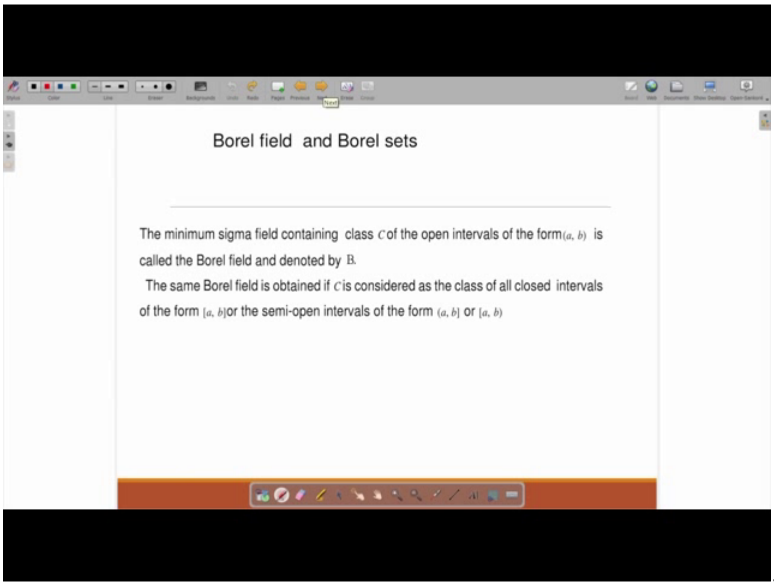
Now, whether close intervals a b belongs to F, we can see that close interval a b can be expressed as the compliment of that is union of the open intervals minus infinity to a b to infinity and then if we take the complementation these are all sigma algebra operation therefore, the result that is the close interval a b that is a member of F.

Similarly minus infinity b this is a semi infinite semi open interval and these we can write as the union countable union n is equal to 1 to infinity of the sets that is we can consider it close sets minus n, b. Now if I consider the union countable union of all the sets of this form then we will get this semi infinite semi close interval minus infinity to b. Therefore, this also belong to F.

Similarly, we can show that this b be the single ton point single point set so, single ton set. So, whether this point sets R a member of this collection this we can consider as suppose if I consider b minus 1 by n, n b this close interval if we consider and then you do the intersection countable intersection from n is equal to 1 to infinity, then we will get only one element that is b therefore, this b also belongs to the sigma field F.

Similarly, we can consider we can show that other interval like b to infinity this will be a member of the field F. So, in this way we can generate a minimum sigma field which starting with a all open intervals and this sigma field will be called the Borel field, borel sigma field.

(Refer Slide Time: 13:38)



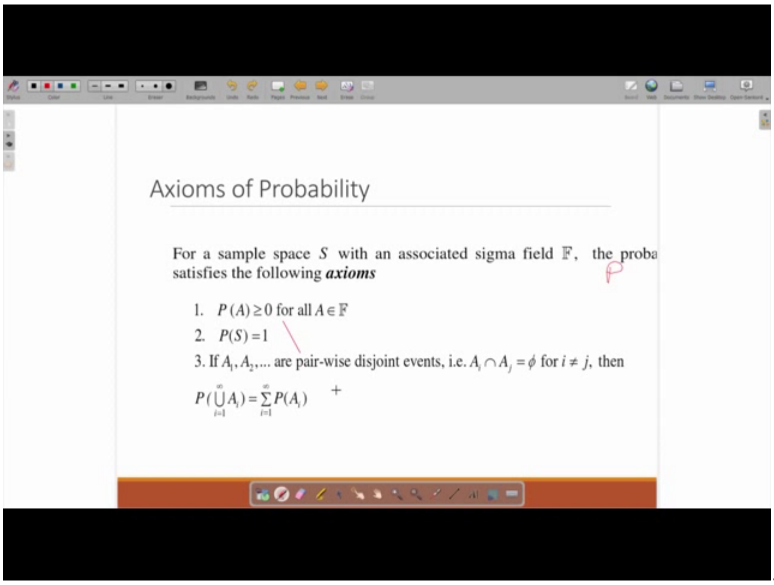
So, what is a borel field and a borel sets? The minimum sigma field containing class C of all open intervals of the form a, b is called the borel field and it is denoted by letter B.

So, we showed in the earlier example how to generate this sigma field and this sigma field is known as the borel sigma field and all the members are called borel sets.

The same borel field is obtained if C is considered as the class of all closed intervals instead of open interval a, b we can consider close interval a, b of the form a, b that is a close interval or the semi open interval of the form one side closed other side open or this left side closed right side open etcetera. So, that way we can define the borel field and this borel field is needed to define probability on the real line and this can be easily extended to higher dimension.

Now, we will formally define probability.

(Refer Slide Time: 14:53)



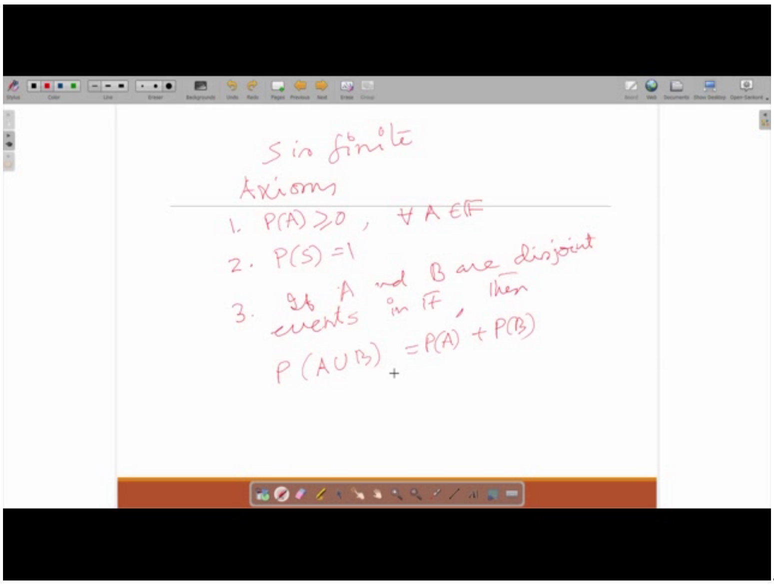
Probability is defined axiomatically on the so, for a sample space S with an associated sigma field F the probability P, probability P satisfies the following probability P satisfy the following axioms. Number 1, probability is a function on set therefore, this P of A that is the probability of event A it is always greater than equal to 0 for all A belonging to F, probability is a non negative function. Then probability of S is equal to 1. So, for the entire sample space probability must be equal to 1.

Now, third axiom that is important if A 1, A 2 etcetera are pair-wise disjoint events that is for i not equal to j, A i intersection A j equal to phi, then the probability of the countable union of this joint events this is the probability of the countable union of disjoint events

A i s. So, this probability is equal to the summation of the individual probability. So, if we have a collection of disjoint events they are union probability is equal to the sum of the individual probability this is known as the sigma additivity property. So, with these 3 actions now we can derive all other properties of probability and any other definition of probability also must satisfy this 3 axioms.

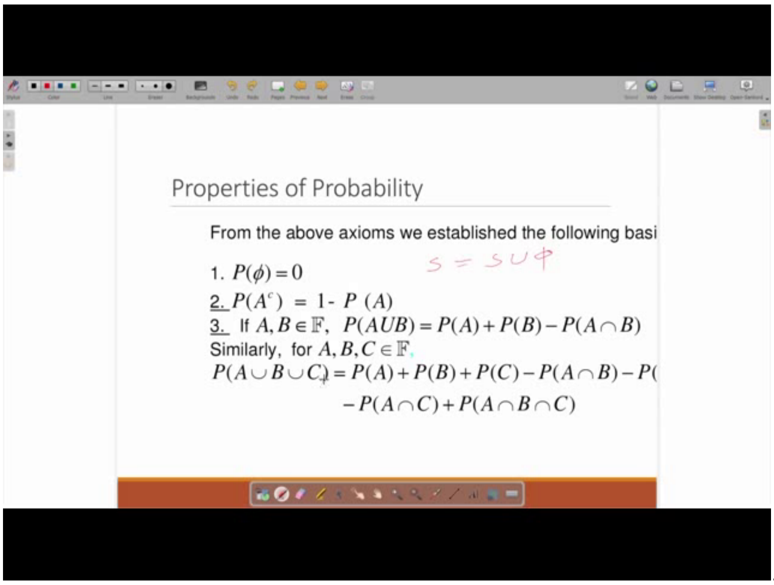
Now, for a finite sample space the axioms particularly the third axiom can be simplified. So, first 2 axioms are same so, if S is finite.

(Refer Slide Time: 16:56)



Suppose S is finite, in that case axioms are what are the axioms number 1, probability of a is always greater than equal to 0 for all A belonging to F that is number 1, it is same as the earlier axiom we have considered. Similarly number 2 probability of S is equal to 1, then third axiom can be simplified and that is if A and B are disjoint event, events in F then probability of A union B is equal to P A plus P B. So, this is the axiomatic definition of probability defined on a finite sample space S.

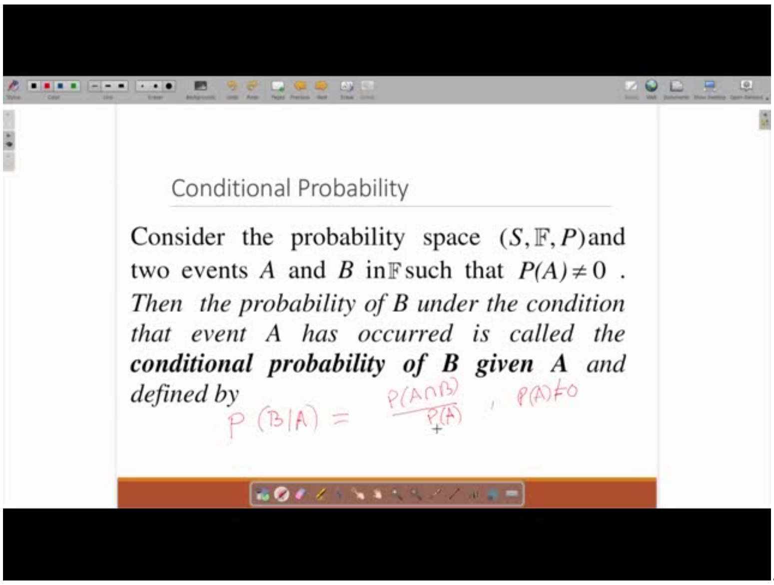
(Refer Slide Time: 18:03)



Now, other properties of probabilities can be derived using these axioms for example, probability of S is equal to 1. Therefore, we know that S is equal to S union phi therefore, probability of S is equal to probability of S plus probability of phi. So, from which we will show that probability of phi is equal to 0. Similarly probability of A compliment is equal to 1 minus P A.

If now if A and B belongs to F then need not be designed then probability of A union B is equal to probability of A plus probability of B minus probability of A intersection B. Similarly if 3 events are there A B C belongs to, F then the probability of union will be now, all individual probabilities will be added then joint probability will be separated and that probability of joint probability of A B C will be added, this principle can be extended to consider the probability of union of any number of events.

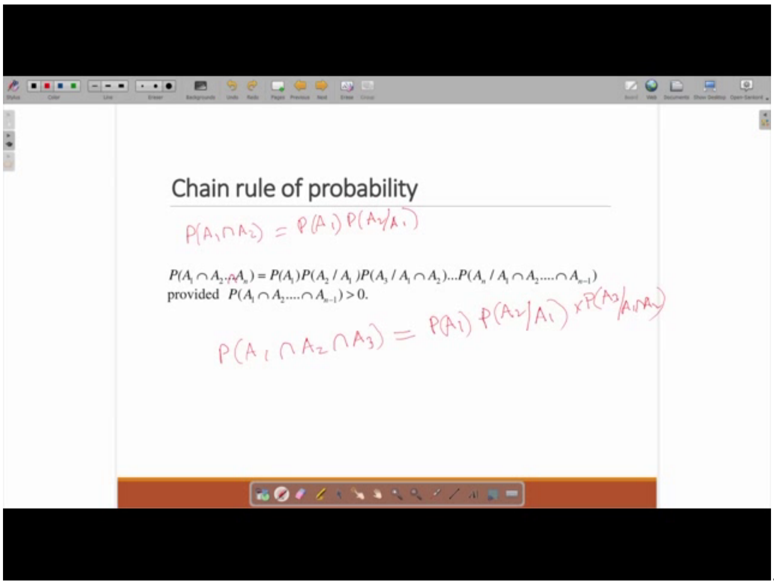
(Refer Slide Time: 19:28)



So, that way we define probability using at 3 axioms and then also we showed that the probabilities we define the probabilities using 3 axioms and we showed that the other properties of probabilities can be derived using these 3 axioms. Next we will consider the conditional probability, conditional probability first we have to define and this definition should be in consistence with the axiomatic definition of probability.

Consider a probability space S F P this triplet sample space S sigma field F and the probability P and 2 events A and B in F such that P A is not equal to 0, then the probability of B under the condition that A has occurred given that A has occurred what is the probability of B. So, this we write as probability of B given A. So, by definition this will be probability of A intersection B divided by probability of A. So, this is the definition probability of A not equal to 0. So, we have to find out the joint probability of A intersection B divided by probability of A that is the conditional probability.

(Refer Slide Time: 21:11)

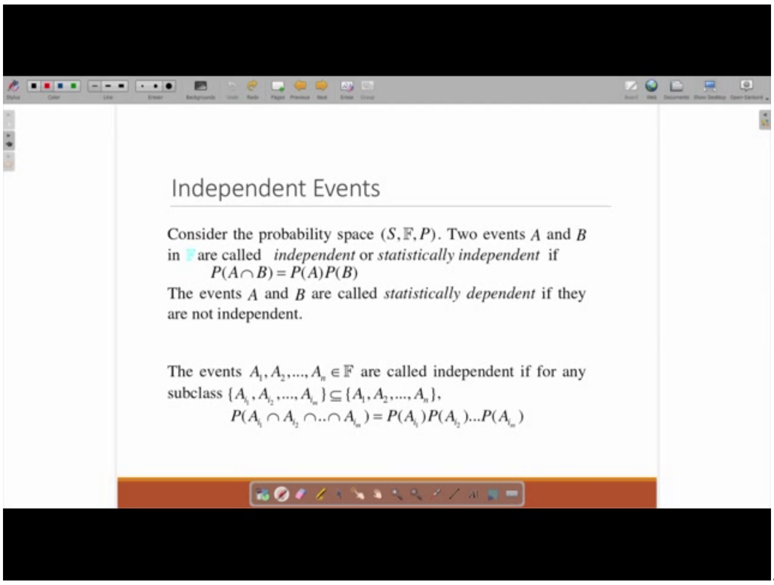


.

Now, given the conditional probability we can find out the joint probability for example, we can write probability of A 1 intersection A 2 that is equal to probability of A 1 into probability of probability of A 2 given A 1 this can be obtained from the definition of conditional probability itself. So, using this definition we can find out the joint probability of n events A 1, A 2 up to A n the joint probability you can find out what is this that is probability of A 1 into probability of A 2 given A 1. In the probability of A 3 the third event we will consider given A 1 intersection A 2, similarly when we consider the last event A n we have to consider all other events A 1, A 2 up to A n minus 1. So, this rule is known as the chain rule of probability using this probability using this chain rule we can find out the probability of the intersection of events.

For example probability of A 1 intersection A 2 intersection A 3, these 3 events their intersection probability will be given by probability of A 1 this is the first event, in the probability of second event given A 1 given the first event into probability of the third event A 3 given A 1 intersection A 2. So, that way we can define we can find out the joint probability of 3 events, 4 events etcetera using this chain rule of probability. So, we have defined the conditional probability and from using the conditional probability we showed how we can find out the joint probabilities.

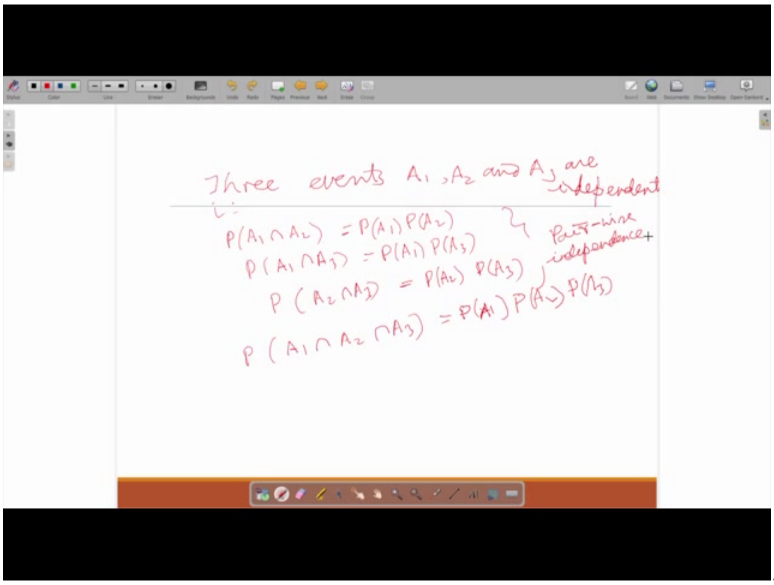
(Refer Slide Time: 23:23)



Next we shall consider independent events. So, 2 events A and B are called independent or statistically independent sometimes it is called so, if P of A intersection B this is the joint probability is equal to probability of A to probability of B. If this relationship is not true, then we call that events A and B are statistically dependent that way we define the independence of 2 events how do I define independence of multiple events.

The events suppose there are events they are independent if we consider any sub collection of these events suppose A i 1, A i 2 to A i m we are considering and this is a subset of the entire collection. If this joint probability of this intersection of these subsets is same as the product of the individual probabilities, then we say that this event and events are independent. So, we are considering for any subclass if the joint probability is product of the individual probabilities than these events A 1, A 2 up to A n events are independent we can consider an example with the 3 events, A 1, A 2 and A 3.

(Refer Slide Time: 24:54)

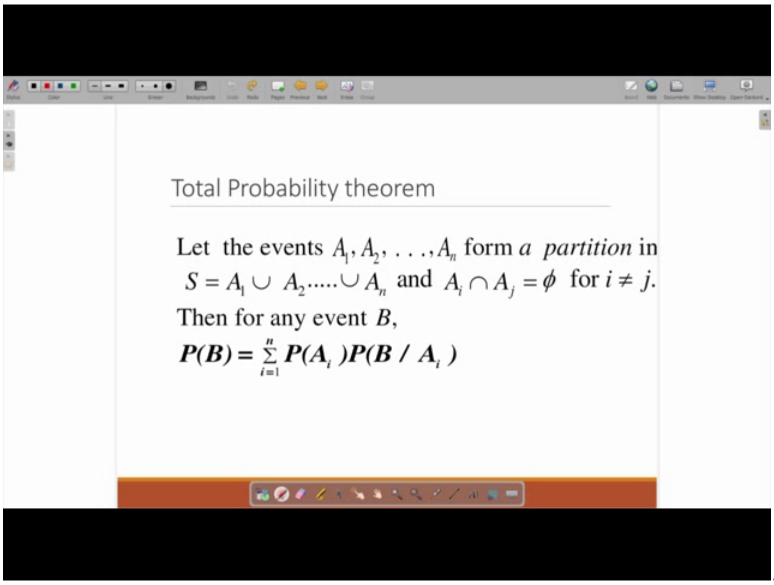


So, when are they independent so, they are independent if probability of first we have to consider A 1 intersection A 2 is equal to probability of A 1 into probability of A 2, probability of A 1 intersection A 3 that is equal to probability of A 1 into probability of A 3, probability of A 2 intersection A 3 is equal to probability of A 2 into probability of A 3 and finally, probability of A 1 intersection A 2 intersection A 3 is equal to probability of A 1 A 1 into probability of A 2 into probability of A 3.

So, if all these 4 conditions are satisfied, then we say that 3 events A and B are independent are independent if this condition is these 4 conditions are satisfied, sometimes these 3 conditions are called pair wise independence and these conditions this condition. So, for 3 events A and A 1, A 2 and A 3 to be independent all these 4 conditions are to be satisfied, if only first 3 conditions are satisfied then the events A 1, A 2 and A 3 will be called pair wise independent. These we will keep pair wise pair wise independence. So, we have introduced the concept of independence.

Next we will go to one important concept that is total probability theorem.

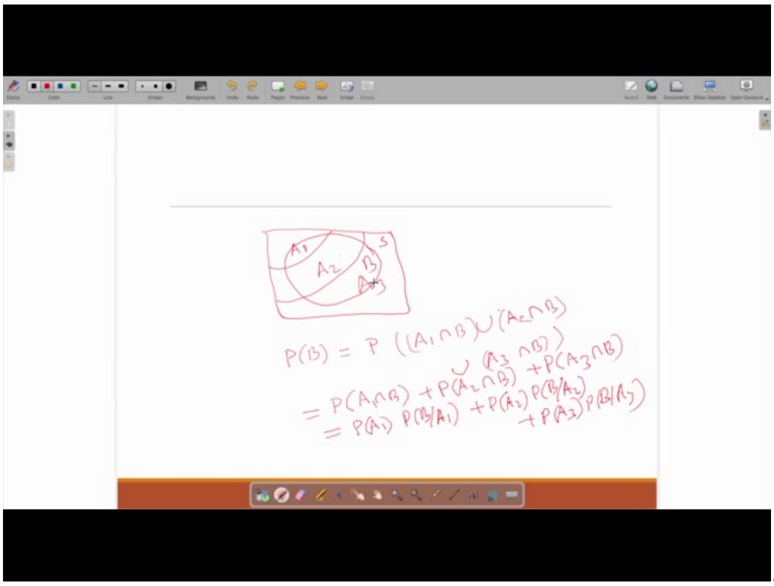
(Refer Slide Time: 27:23)



So, this theorem says that if events A 1, A 2 up to A n if they form a partition in S what that does it mean that S is the union of these events and all these events are mutually designed. Then probability of B is equal to this will be summation i is equal to 1 to n of probability of A i into probability given A i.

So, if we consider the product of P of A i into probability of B given A i and sum of for i is equal to 1 to n, then we will get the probability of B this is known as the total probability theorem we can illustrate this with the help of a simple case.

(Refer Slide Time: 28:18)



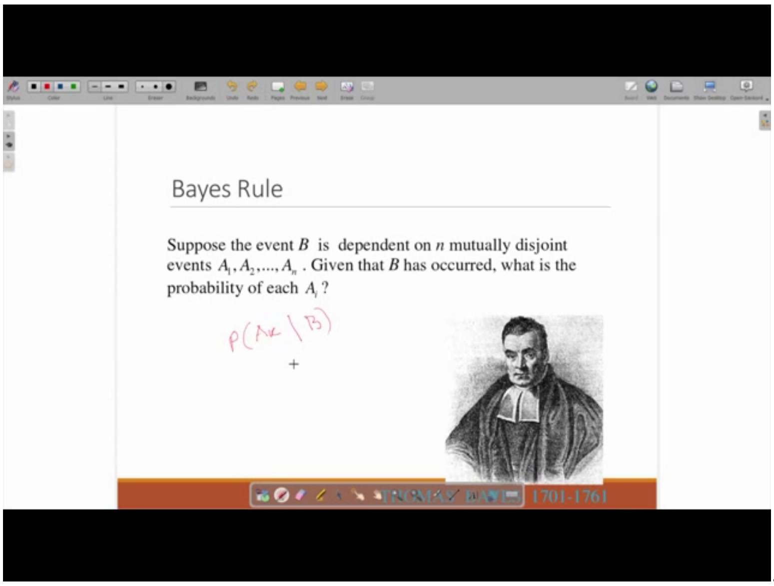
Suppose we have this is my S and we are considering a partition A 1, A 2 and A 3 and B is an event like this is my B.

So, what I see that B is the union of this part, this part, and this part therefore, probability of B will be equal to probability of A 1 intersection B union A 2 intersection B union A 3 intersection B.

Now, all this intersections A 1 intersection B A 2 intersection B this is A 1 intersection B this is A 2 intersection B this is A 3 intersection B they are designed therefore, the probability of the union will be equal to the sum of the probability that is equal to probability of A 1 intersection B plus probability of A 2 intersection B plus probability of A 3 intersection B.

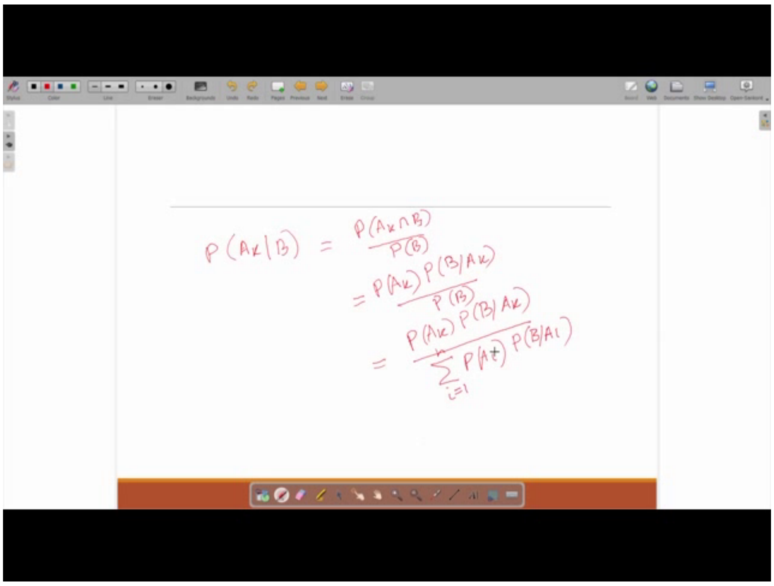
So, these now we can apply this joint probability as the product of individual probability and the conditional probability therefore, this will be probability of A 1 into probability of B given A 1 plus probability of A 2 into probability of B given A 2 plus probability of A 3 into probability of B given A 3. So, this is the total probability theorem in the case of 3 A partition of 3 events A 1, A 2, A 3 etcetera this can be the for the any number of events.

(Refer Slide Time: 30:30)



Now, using the total probability theorem and the condition definition of conditional probability we can derive the Baye's rule. Suppose event B is dependent on n mutually disjoint events A 1, A 2, A n, given that B has occurred what is the probability of each A i? So, suppose B has occurred we are interested to find the probability of any A k you consider given B. So, how we can find out?

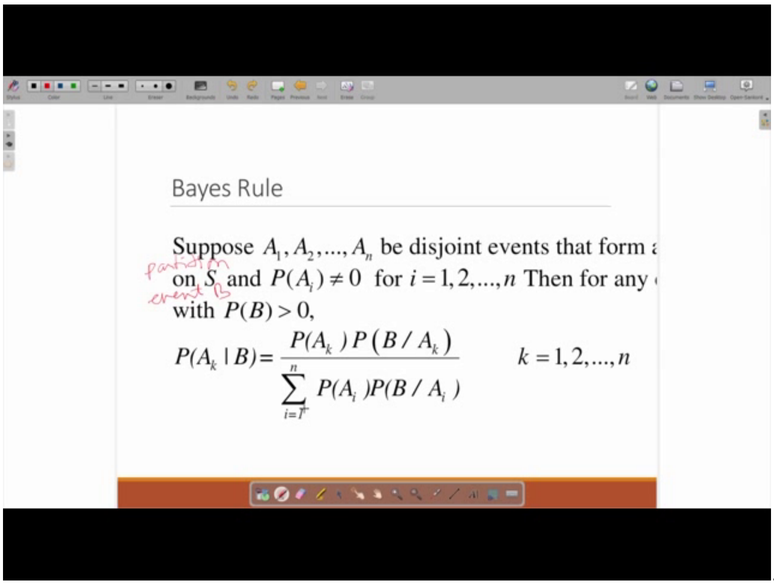
(Refer Slide Time: 31:06)



Now, probability of A k given B so, this is equal to probability of A k intersection B divided by probability of B this is by definition of conditional probability. Now, if we write this S probability of A k into probability of B given A k divided by probability of B. Now we have to find out what is the probability of B that we can find out using the total probability theorem. So, this will be given by a summation that I will show you.

So, what is probability of B, probability of B is nothing, but summation i is equal to 1 2 n probability of A i into probability of B given A i. So, that way we can find out suppose the event B has occurred what is probability of a particular A K. So, that we can find out using this result what is known as the Baye's rule, this is due to a famous mathematician aa named Thomas Baye.

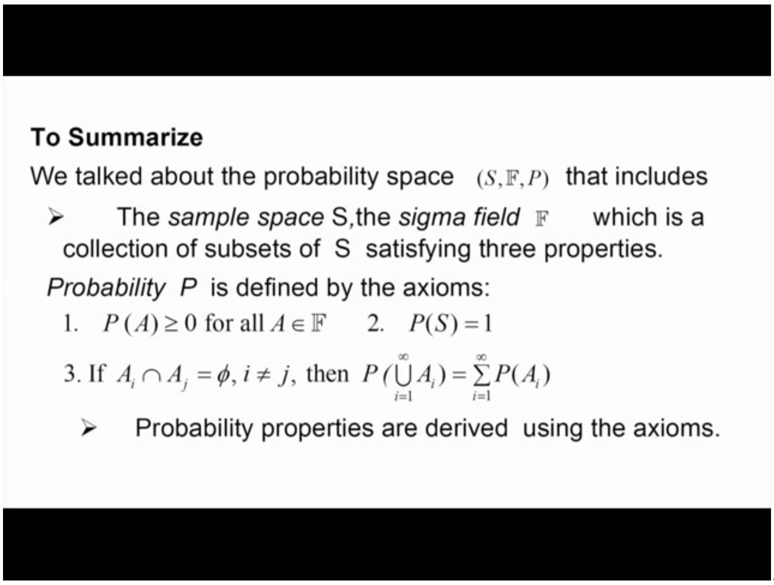
(Refer Slide Time: 32:33)



So, we can formally said that let us see so, this is the Baye's rule, now formally we set this suppose A 1, A 2, A n be disjoint events that form a partition form a partition this is a partition on S and probability of A i not equal to 0, for i is equal to 1, 2,.. n. So, all these probabilities are known suppose probability of A i this event A i that is they are known and none of them is equal to 0.

Then for any event B with probability P B greater than 0 now what we will get probability of A k is given by this base formula this is the Baye's rule, now using this Baye's rule we can find out suppose given some evidence what is the probability of a particular cause. So, this is very important of when we apply probabilistic decision.

(Refer Slide Time: 33:46)

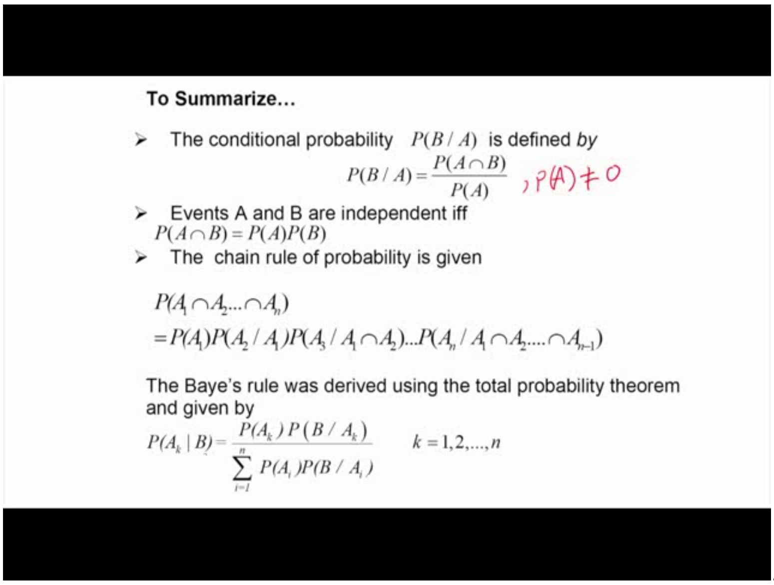


Let us summarize the lecture we talked about the probability space S, F, P that includes the sample space S, the sigma field F, now F is a collection of subsets of S satisfying 3 properties. The 3 properties were that S is a member of F, S is closed under F is closed under complementation and F is closed under countable union and probability P is defined using 3 axioms.

These 3 axioms are probability of A is greater than equal to 0 for all A belonging to F that is for all events in F probability of A must be greater than equal to 0, probability of this sample space is equal to 1. Then third axioms was if A i intersection is equal to phi; that means, the events are designed then probability of the countable union of A i is equal to sum of the corresponding probabilities.

So, these 3 axioms were used to derive other properties of probability. So, these axioms were used to derive other properties of probability.

(Refer Slide Time: 35:18)



We also define the conditional probability P of B given A, this is defined by this relationship probability of B given A is equal to probability of A intersection B eh divided by probability of A of course, P of A not equal to 0. Similarly we could define probability of A given B, then events A and B are independent if probability of A intersection B is equal to P A into P B.

To find the joint probability of any number of events we use this chain rule. So, according to chain rule probability of A 1 intersection A 2 up to intersection of A n that is given by probability of A 1 probability of A 2 given A 1 probability of A 3 given A 1 intersection A 2 like that up to probability of A n given A 1 intersection A 2 up to A n minus 1. So, this is the chain rule of probability then we derive the aa Baye's rule using the total probability theorem.

So, it finds the A (Refer Time: 36:41) probability, probability of A k given B that is given by probability of A k into probability of B given A k divided by probability of B actually if I apply the definition of conditional probability then this denominator should be probability of B and this probability of B according to the total probability theorem is given by this summation. Now here probability of A k these are the probability and what we are determining that probability of A k given B that is the A (Refer Time: 37:23) probability.

Thank you.