Advanced Engineering Mathematics

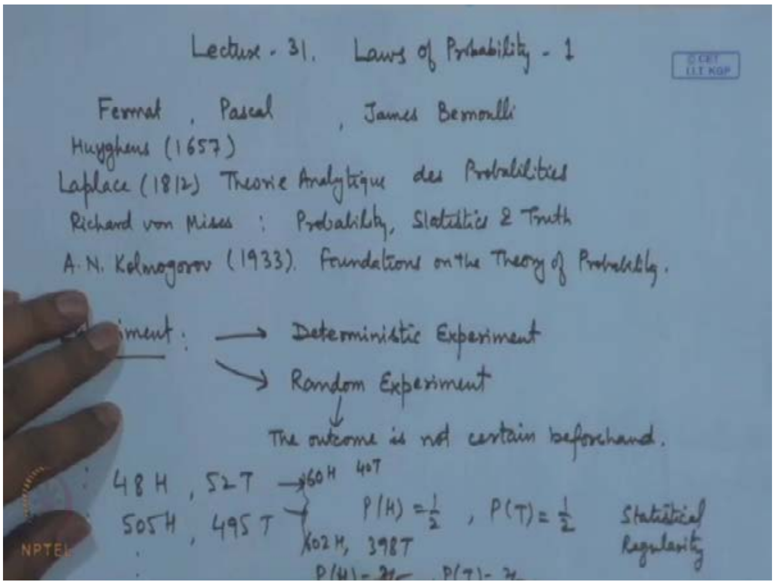
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture No. # 32
Laws of Probability - I

In this particular module, we will we will be discussing fundamentals of probability, and statics. Nowadays probability, and statics plays a very important rule in all walks of life, all engineering, and science disciplines in social sciences. In fact, we can say that there is no area of human activity, where probability or statics is not used. The subject of probability as old as the civilization itself; however, the modern theory of probability as we know today as its routes in games of chance, and particularly in 15th and 16th century Europe, when the (( )) games we are been played (( )) in the form of dice games, cards games, etcetera.

Then some of the people got interested in knowing the that what are the chances, that if the bit on a certain event, then whether it will be more likely or less likely. For example, in a grame of cards, they would like to know whether a particular player will get all the 4 asses or all the for kings or whether you will get the top 4 cards of a particular denomination.

(Refer Slide Time: 02:03)



This let us to the certain correspondence, and activity among the mathematician especially 2 of the... So, today we will discuss basically the laws of probability - the basic laws of probability. The some of the famous mathematicians of the day for example, Fermat Pascal, they had famous correspondence in which we discuss the elementary laws of probability. However, the first publish to our is probably by Huygens in 19 in the year 1657, in his book; there is also contribution by James Bernoulli around the same time.

The landmark are you can say the first milestone in the development of the subject probability can be considered by the monumental work by Laplace, the French mathematician in his book theory analytique des probabilities Laplace covered all the development of the subject of probability, which was known till that day. So, that he added is a own theories, and this gave a foreign foundation and a mathematical treatment to the subject probability.

Some of the famous contribution to the subject of probability are by von mises probability is statistics, and truth. The model probability theory that is after the (( )) development of that any mathematical theory should have an arithmetic setup, and from there the entire theory should be able to derived. This let to the arithmetic development of the probability theory, and A N Kolmogorov the Russian mathematician published a book in 1933, foundations on the theory of probability.

(No audio from 04:35 to 04:49)

Let us see that, what are the fundamental units that wants should know or the fundamental definitions for the development of the subject of probability. So, when we give a typically statement, that it is likely that it will rain today or I may miss by train today or my for example, if I am doing a study, I may say that my study is likely to be successful or unsuccessful. This type of statements or in a sense giving as a measurement of how much likely on event is, but in the first place we may not actually give number, we may just say that it is likely that it will rain today without quantifying it. However, if we say that there is a 75 percent chance that it will rain today, then we are putting measurements with that.

So, the subject of probability is related to giving numerical values to the probabilities of various statements. And therefore, it deals with certain phenomenon, where things are uncertain. For example, let us look at two types of experiments: One experiment is that we take to molecules of a hydrogen, and molecule of a oxygen and we mix them. We know that the reaction of it will be lead to the water. Now, this is an experimented chemistry, and similarly there are various experiments in science and engineering, where we carry out of experimental under certain conditions, and we come with this certain conclusion. We know that this will be the outcome of the experiment, such experiments are known as deterministic experiments. So, fundamental unit you can say in the study of the subject probability, and statistics is an experiment.

So, an experiment can be a deterministic experiment, I give the example of a deterministic experiment just now. Another experiment could be where, if we contact the experiment we are not sure of what outcome will be there. For example, if we task a queen, then we may get a head or tail or we may also include the possibility that it may stand on its side. If we toss a dice, then we may get any face up for example 1, 2, 3, 4, 5, 6, if we talk about the weather tomorrow, then the weather could be sunny, it could be cloudy, it could be rainy, there could be thunderstorm.

If we are talking about say next 10 years, say it states of the number of earthquakes in a particular seismic zone, then there may be one earthquake, there may be no earthquake, there or there may be 3 earthquakes or there may be 15 earthquakes and so on. Now, these are the kind of events, where we are unable to know the outcome of the experiment before and, now we say experiment - experiment does not mean that only laboratory experiment; experiment means observing or contacting something under certain conditions. So, when we say whether then we are only observing the weather, because as we know that there will be is spring, there will be some are there will be raining season, then there will be winter and so on.

So, the weathers happen due to certain natural phenomena; however, when we want to observe that then it is considered as an experiment. And if we are looking at say day today or behaviour over a period of time, then it is a random experiment. A non deterministic experiment is known as a random experiment, because we are not sure of what outcomes will be there in the beginning. So, random experiments are the one, here the outcome is not certain beforehand. Now, one may just question that if we are not knowing the outcome before, and then why do we actually study the subject of probability.

For example if we are looking at the tossing of a queen, then suddenly it can be head or tail, but it is not known at each toss of coin, whether we will get head or tail. But there is other feature of such experiments, which allows as to study the subject of probability, and a more theoretical basics. For example, if we contact the experiment say 100 times, suppose the coin is the fear coin, we may observe something like say 48 heads, and say 52 tails; out of 100 times.

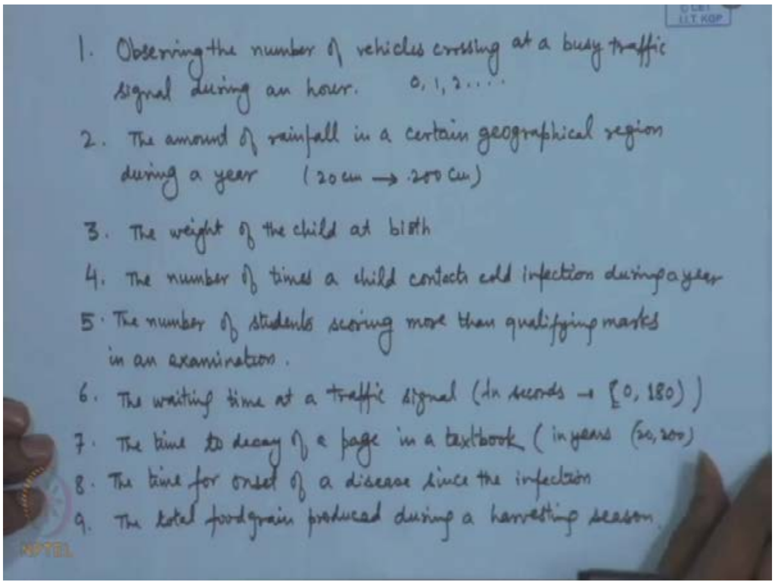
Suppose, we conduct the experiment of tossing of the coin 1000 times, we may observe the there are say 505 heads, and say 495 tails. That means, if we are contacting at a large number of times, we may observe a pattern; and that pattern gives as the probability of individual events. For example, we may safely say that the probability of occurrence of a head is half here or the probability of the occurrence of tail is half here. Similarly, if we today occur that my contacting of this experiment would have l yielded. For example, in the toss of 100 times, suppose we have 60 heads, and say 40 tails.

Similarly, suppose in 1000 we may get something like say 602 heads, and say 398 tails. Then we may be more and clime to say that it say bias coin in the favour of head, and we may put say probability of head as say 2 by 5 sorry 3 by 5, and probability of tail as 2 by 5. Therefore, what we observe here is that individual experimental outcomes are not known, but when, but there is a long term its statistical regulatory; this term is known as its statistical regulatory. That is the long term behaviour that one can say.

So, for example, in a 1000 births in a city hospital in a period of say 6 months. We may except that there are 500 girl child's, and 500 boy child's born. If we observe the weather pattern over several years in a monsoon region, then we may say that during the amount of rain fall may be say 100 centimetres or if we are observing the weather pattern of drought and rains, we may say that after every 10 years there is a likelihood of a drought year. So, for example, we may say that the probability of a drought is1 by 10.

So, this long term behaviour that we can make out, that we can predict out a uncertain event; justifies the a study of the subject probability. Now, we introduce the examples of random experiments. So, we may consider various kinds of phenomena. For example, I just mentioned about say tossing of a coin throwing of a die, ignore of a car form a (( )) car etcetera.

(Refer Slide Time: 13:12)



However, these are not only text book kind of example that are random experiments, there can be many more. For example, we may observe observing the number of vehicles crossing at a busy traffic signal during an hour. So, here for example, the number of the vehicles could be 0, 1, 2, and so on. The amount of rainfall in a certain say geographical region during a year. So, depends upon the region suppose the region is play a such that, this is lot of rain then this amount could vary from say 20 centimetre onwards to say 200 centimetre in have interval like this. The weight of the child is recorded at the birth, so different children will have different weight at birth. So, then a vary from a few 100 grams to few kilograms kind of thing. The number of times a child contacts a cold infection during a year. The number of students scoring more than qualifying marks in an examination.

So, for example, the total marks are 100, and the qualifying marks are fixed as a 60; therefore, how many students out of the total number of students which are appearing, how many of them will cross that qualifying marks that will become a random experiment. You are going on a route and you are a traffic signal, so the traffic signal may be green. So, you may just cross or it may be red in that case you may have to wait depended upon the total duration of the signal. For example, it may vary from 0 to 3 minutes.

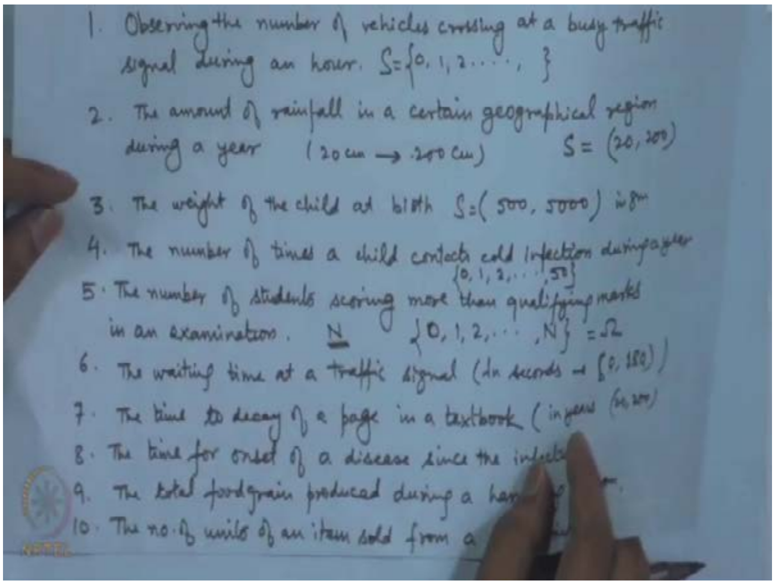
So, the waiting time at a traffic signal, so suppose we are according in seconds, then it could be vary from 0 to180 second; the maximum limit till which the signal may be a having a particular sign. The time to decay of a page in a text book. So, this could be the time recorded in years; for example, it may take say 20 years to 100 years for the periods to decay or may be 20 years to 200 years for the page to decay. The time for onset of a disease since the infection, so for example, one is infected with say HIV virus, and then the time when the aids disease develops in the person. So, that is a random experiment.

The total food grain produced during a harvesting season. The number of units of an item sold from a store during a day. So, you can see that the examples of random phenomena as very yield as possible, and the examples are ranging from engineering physics, medical, social sciences, economics.

So, almost there is no area human activity, where the random phenomena is not there. In fact, the quantum mechanics assume that the moment of the electrons is random. And that is why we the modern theory of physics are you as it called as statistical physics is there.

Now, we look at some basic terminology which is used in the subject of probability, and then we will find the probability. So, let me take a basic unit of a random experiment is a sample space. So, a sample space is the set of all possible outcomes of a random experiment. So, the usual notations one can be use S, omega, theta, etcetera. So, let me look at the example that we discuss just now.

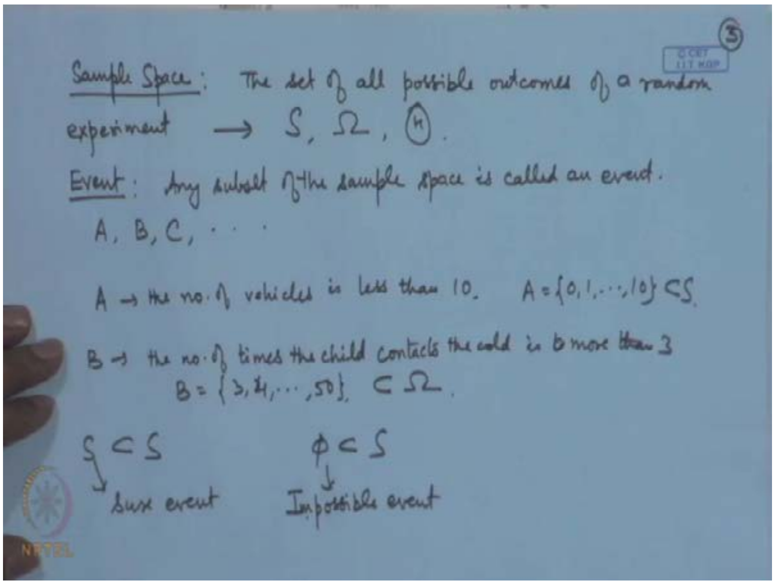
(Refer Slide Time: 20:22)



So, for example, if we are looking at the number vehicles crossing at a busy traffic signal during an hour. And here the sample space we may write as the numbers 0, 1, 2, and so on. If we are looking at the amount of rainfall in a certain geographical region during a year, and as I mention that the area is such that it receives lot of rain, then my sample space can be expressed as a an interval 22, 200 where the unit of measurement use in centimetres. The weight of the child at birth, so it may be as it small as a few hundred grams to say 500 grams to say 5000 grams; it may vary little more also depending upon what features we are study here. The number of times a child contact cold infection during a year; once again the number could be 1, 2, and so on. And it may end up at a finite, because the total number days in year is 365. So, suddenly the number cannot be very large, you may put say 0, 1 to 50. The number of a student is scoring more than qualifying marks in an examination.

So, suppose the total number of students are N, which are taking the exam, then the sample space could be the waiting time at a traffic signal in seconds for example, it would be 0 to 180, the time for decay. So, in un experiment one, define what is the sample space depended upon what is our area interested.

(Refer Slide Time: 22:08)

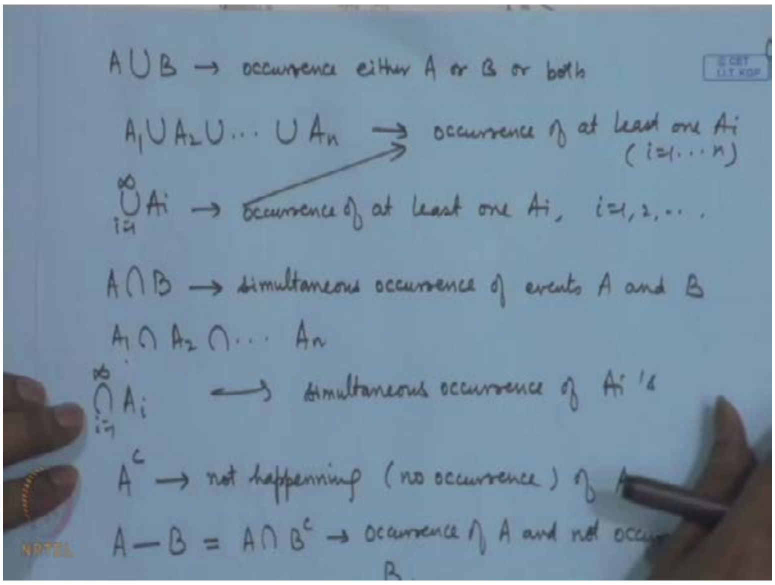


Then, what is an event? Any subset of the sample space is called an event. And we usually imply English letters in capital to denote the events. So for example, if we look at the number vehicles crossing at a busy traffic signal. I may define the event A as the number of vehicles is less than 10; the naturally A is consisting of 0, 1, 2 10. This is the subset of S. I may consider say the number of times a child contacts cold. So, you may say the number of times the child contacts the cold.

So, we may a put is say more than 3; in that case the set B will be 3, 4 and so on. Suppose, we are putting the upper bound (( )). So, this is subset of… Here if I it denoted by say omega, then this is the subset of omega. So, event is the subset of the sample space; however, we may have extreme cases, then we say subset, then empty set is also a subset, the full set is also a subset. So, if we say full set. So, S is a subset. So, this is called sure event, and phi is the subset of S this is known as impossible event.

For example, if we say that the weight of a child at birth minus 32, then it is impossible event. Suppose, we say that the number of vehicles crossing at a busy traffic crossing is say one million, then there is will be a impossible event, because it cannot cross the total number of vehicles which are available there, and so on. Now, when we associate sets with the events, then there are set operations like union, intersection, complementation, etcetera. So, in terms of a events they have various interpretations I would like to explain this now.

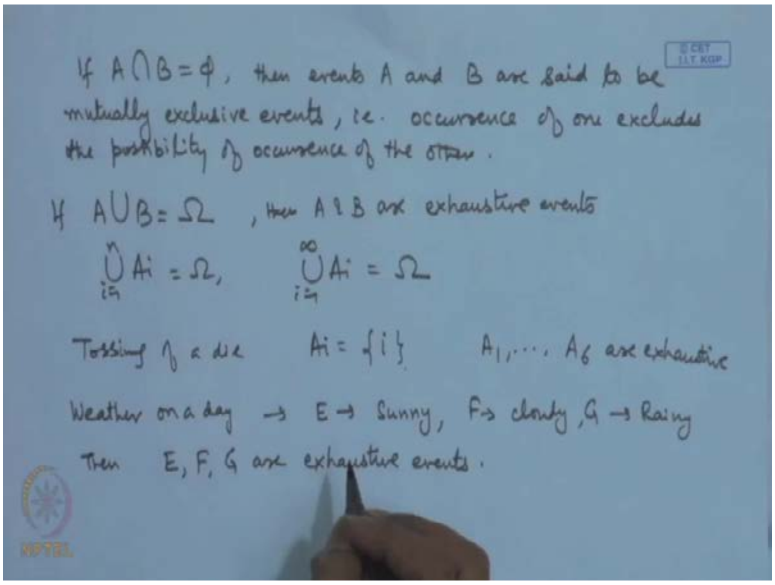
(Refer Slide Time: 25:05)



So, we may consider say A union B; now if there are 2 events A and B, A union B represents occurrence of either A or B or both. Now, we can generalize notation for example, if I have A1 union A 2 union A n, then we may say this is occurrence of at least one A i for i is equal to1 to n. We may even talk about, and in finite union i is equal to1 to infinity; the interpretation of this will also the same, occurrence of at least one A i, occurrence of at least one A i; now here i will be 1, 2, and so on.

Similarly, if we consider the concept of intersection of sets, then here intersection of the sets will denote the common elements belonging to A and B. So, this means that both A and B occur. So, this we can say simultaneous occurrence of events A, and B. Now, this concepts can further we generalized to n events or infinite number of events also. So, this is simultaneous occurrence of A i, that is all A i's occur together. Then there is a concept of complementation for an event A - A complement denotes not happening or no occurrence of A; similarly we may interpret A minus B. A minus B in the set theory denotes the set of elements which are in A, but not in B. So, this will become A intersection B complement; that means, occurrence of A, and not occurrence of B. Then there is a concept of (( )), because when we consider the concept of intersection, the intersection could be phi also.

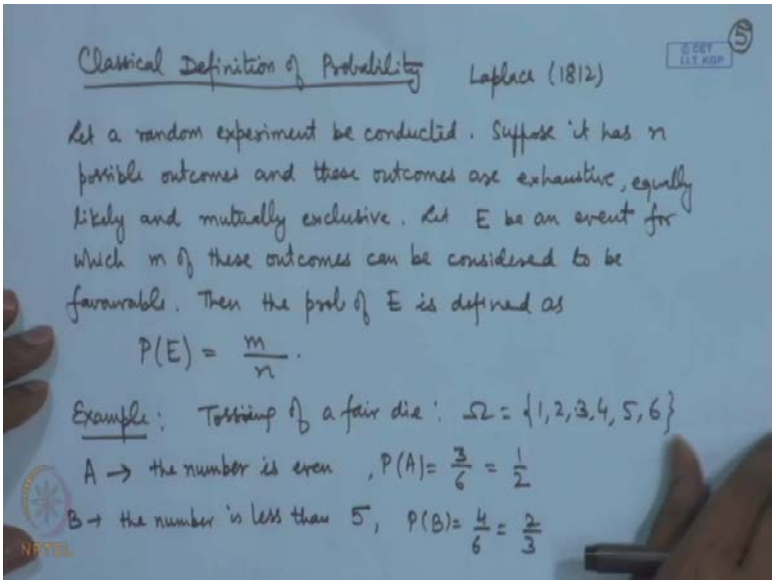
(Refer Slide Time: 28:03)



If the intersection is phi; if A intersection B is phi, then they are known as disjoint set. So, here we call them mutually exclusive events; the meaning is that if A occurs B cannot occur, and if B occurs then A cannot occur. So, this is events A and B are said to be mutually exclusive events; that is occurrence of one excludes the possibility of occurrence of the other. There is also a chance that some of the events for example, A union B is equal to the full sample space; if A union B is equal to omega, then all the possibility of the sample space are considered by A and B, then we say A and B are exhaustive events.

This can be generalized, we may have i is equal to1 to n is equal to omega or we may say union A i, i is equal to 1 to infinity is equal to omega. See for example, if we are considering tossing of a die, if we are tossing of a die, and we consider A i as i. Then A1, A 2, A 6; they will exhaust all the possibilities of 1 to 6. So, they are exhaustive, if we are looking at say weather on a day, and we define the events say E as sunny, F as say cloudy, and G as say rainy. Then all the possibilities of the weather are exhausted, and we may say E, F, G are exhaustive.

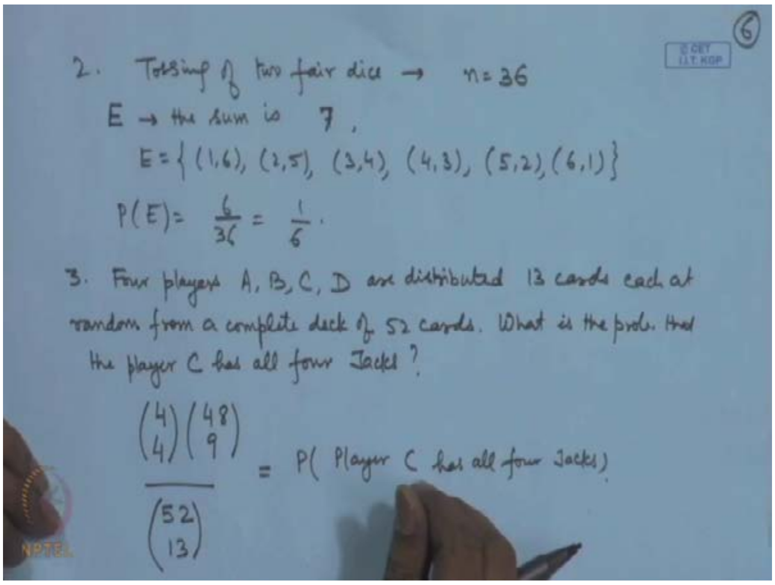
(Refer Slide Time: 31:00)



Now, we are ready to look at the one of the preliminary definitions of probability, which we call as a classical definition of probability. This can be attributed to Laplace; he was the first one who gave it in this particular form. Let a random experiment be contacted; suppose it has n possible outcomes, and these outcomes are exhaustive. That means, we have consider all the possibilities, equally likely – equally likely means that each of them has the same chance of a appearing, and mutually exclusive. That means, occurrence of one will be excluding the possibility of the occurrence of the other. Let E be an event for which m of these outcomes can be considered to be favourable. Then the probability of E is defined as probability of E is equal to m by n.

So, that definition is as you can see, it is applicable to the experiments, where we have a finite number of outcomes all of which we can an numerate, and we are putting additional restrictions such that they are exhaustive, they are mutually exclusive, and also they are equally likely, in those cases this definition can be applied. Let us look at very simple example say tossing of a fair die, your sample space consist of 6 possibilities, and we associate say event A by saying the number is even. Now, if we want to find out the probability of A, then there are 3 favourable cases 2, 4, and 6; total number of possibility is 6. So, you get half, suppose we say the number is less than 5, then number is less than 5 has 4 possibility here. So, probability of B will be equal to 4 by 6; that is equal to 2 by 3.

(Refer Slide Time: 34:46)



Suppose, we consider say tossing of 2 fair dies, and we say E is the event, that the sum is say 7. Then what are the possibilities here? Then we toss 2 fair dies, the total number of possibility is 36, 1,1,1, 2, 2,1, 2, 2, 2, 2, 6, 3,1, 3, 2, 3, 6, and so on. There will be total 36 possibility, and if we assume that the dies are fair, then each of them will be. So, the set E will be represented by (1,6), (2,5), (3,4), (4,3), (5,2), and (6,1); we are 6 possibilities, which we will lead be the sum 7.

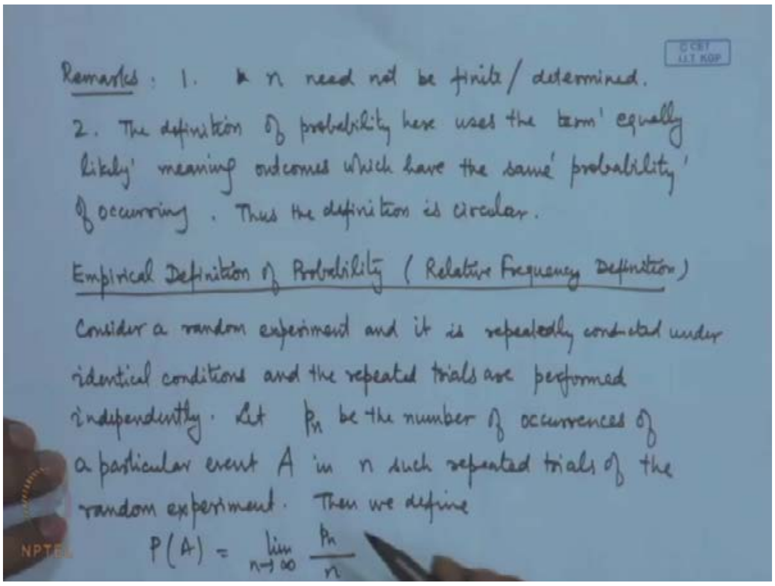
So, probability of E will become equal to 6 by 36, that is equal to1 by 6. Suppose, we take another example here.

(No audio from 36:04 to 36:23)

4 players A, B, C, D are distributed 13 cards each at random from a complete duck of 52 cards. what is the probability, that the players C has all 4 jacks. Now, here we look at this problem, the total number of possibilities for each player. So, there are 4 players, and they are distributed 13 cards each. So, the total number of possibilities for a player C, because we are interested in the event for the player C, the total number of cards that the player C get (( )) 13, out of 52.

So, if we consider the possibility, that it will be 52 C 13. Now, he is getting all the 4 jacks; that means, out of a 13 cards, he is now total there are 52 cards out of its they are 4 jacks, and he gets all the 4. So, 4 C 4 and from the remaining 48 cards, we get any 9 cards. So, this will be the probability of player C has all 4 jacks. So, this answer comes by direct counting; assuming that all the cards are equally likely to be distributed to all the players. This definition is helpful for answering questions of this nature we are the sample spaces are finite, and we are having the facility observing are you can say enumerating being able to enumerate all the possibilities; however, this as some practical difficulties.

(Refer Slide Time: 39:08)



However, this has some practical difficulties; for example, we may have this total number n need not be finite or determined; we may not event able to determined, what is n. Another thing is that we are assuming here that the events are or the outcomes are equally likely; the meaning of equally likely which I mentioned is that they have the same chance of occurrence. Now chance is associated with that our probability; we are defining probability here; in that sense this definition is circular. The definition fails. We are already assuming that each outcome has a equal probability; therefore, this particular type of definition is applicable only to theoretical kind of exercises.

The definition of probability here uses the term equally, likely, meaning outcomes, which have the same probability of occurring. Thus the definition is circular in nature. A more practical definition was developed, and this is based on the empirical evidence. When we made usually statements, it is likely that it may rain today; for example, if we have observed three days of intense heat and humidity in a region, then we say that on the fourth day evening, we may say in the morning, we say today evening it may rain. Now, this is based on our experience. Similarly, when we are observing the performance of the students, and we make a statement, we fix up the qualifying marks 60, and then we say that nearly 50 percent of the students will qualified. Now this is based on our previous experience or previous experimentation for the same event.

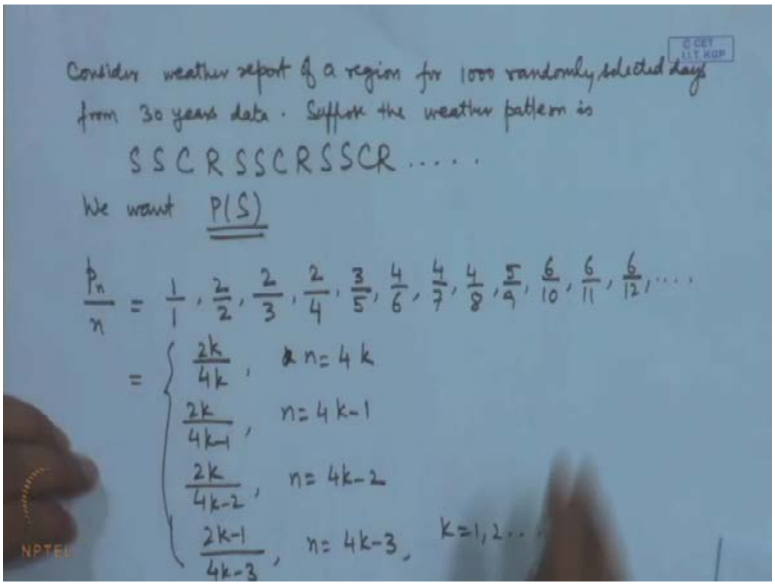
As I already mentioned that the subject probability itself is of interest, because of the feature of a statistical regularity or long term prediction. Therefore, another definition which is a more statistical definition is based on the empirical observations. So, we give empirical definition of probability; this is also called relative frequency definition relative frequency definition. Now, we are looking at the outcomes of a random experiment, based on that certain event is there, for which we are looking at the probability. So, for example, whether it will rain after three days of intense heat etcetera; what is the probability of the students qualifying in a given examination; what is the probability of a patient recovery from a certain disease if he is given a certain meditation.

So, in all these cases, we are observing so for example, patients are being giving certain medicine for certain disease, and then we have the data that how many of them may be recovering, so may be 90 percent are recovering, 80 percent are recovering and so on. Now each unit of observation; so, for example, one patient is being giving the medicine or meditation for a certain disease, then the second patient is given, and this we are observing over a period of time. This is considered as repeatedly contacting the experiment, and we may say roughly that it is contacted under identical conditions.

So, and also we may assume that occurrence are happening of one experiment that means, whatever be the outcome of one experiment does not affect the outcome of the next time. So, for example, one patient is given a certain medicine, he may recover from there; and now a second patient comes with the same disease and the same meditation is given, he may not recover; that means, effect of one occurrence should not be there on the other one. So, we say that the random experiment is repeated under identical conditions, and also independently.

Now, we see how many times over a period of time or how many times over a certain number of trails, this particular event occurs; now this ratio is some value, and over the long range when we are observing over a period of time, this will stabilize; this is known as the probability of that event. So, let me write it here. Consider a random experiment and it is repeatedly contacted under identical conditions and the repeated trails are performed independently. Let say p n be the number of occurrences of a particular event A in n such repeated trails of the random experiment. Then we define the probability of the event A as limit p n by n as n tends to infinity; that means, over the long term, what is the ratio of occurrence of the event, in which we are interested to the total number of trails of the experiment.

(Refer Slide Time: 46:55)

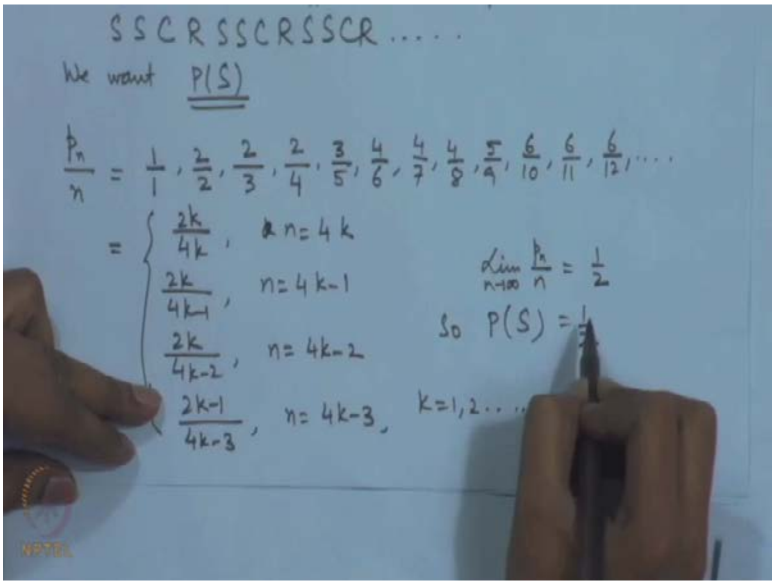


Let me explain through an example here. Consider say let us consider weather report of a region for say 1000 randomly selected days from say 30 years later. Suppose the weather pattern is say sunny day, sunny day, cloudy day, rainy day, sunny day, sunny day, cloudy day, rainy day and so on. The pattern is somewhat fixed that is two sunny days followed by a cloudy day, and then a rainy day. Now we are interested in finding out the probability of, we want the probability of a sunny day. So, we look at this ratio p n by n.

Now this p n by n, when the first trail was there, we observed sunny day, so the ratio becomes 1 by 1. In the 2 trails, 2 days it was sunny, 2 by 2. Now the third turned out to be cloudy. So, the number of sunny day in the 3 trails is also 2. The next turn out to be rainy, so the number of occurrences of the event S was 2 out of 4 also, then 3 by 5, 4 by 6, 4 by 7, 4 by 8, 5 by 9, 6 by 10, 6 by 11, 6 by 12 and so on. Now, we want to find out the limit of this; if we just observed like this, it is difficult to find out the limit, because number in rapidly.

So, we put in a more mathematical form; we can like write like this. It is equal to if we observe each fourth occurrence here, this is of the form 2 by 4, 4 by 8, 6 by 12 and so on; that means, we can write as 2 k by 4 k, whenever n is the form 4 k. If we observe here it is 2 k by 4 k minus1 that is whenever n is of the form 4 k minus1. Here it is equal to 2 k by 4 k minus 2, whenever n is of the form 4 k minus 2; it is of the form 2 k minus1 by 4 k minus 3, whenever n is of the form 4 k minus 3, for k equal to1 to and so on. Clearly you can see here that each of this subsequence converges to half, this is equal to half; as k tends to infinity, this goes to half; as k tends to infinity, this goes to half; as k tends to infinity this goes to half k.

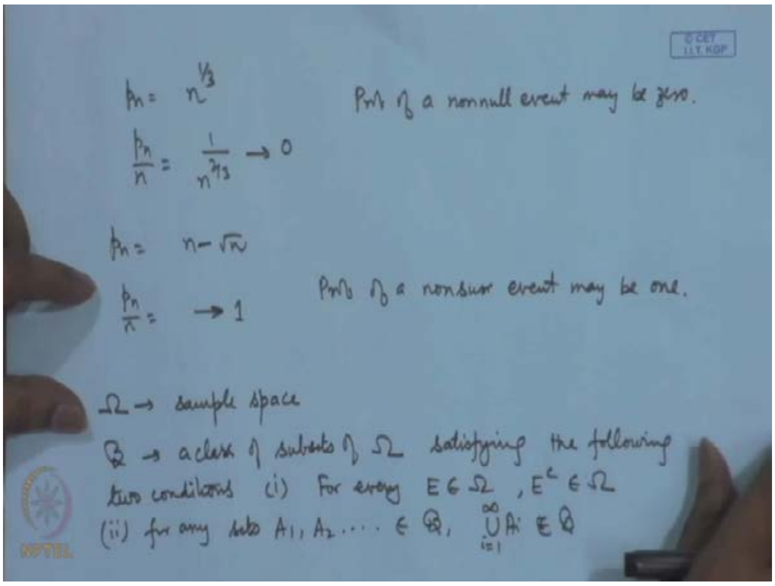
(Refer Slide Time: 46:55)



So, clearly you can say limit of p n by n as n tends to infinity is equal to half. So, probability of a sunny day is equal to half; so this is the experimental, you can say demands station of this method of the relative frequency. So, here you see that 2 sunny days followed by 2 non sunny days, if that is the pattern then definitely the probability of a sunny day over a period of time, should be half. So, relative frequency definition is based on the experience, and this is the most widely applicable definition of the probability today.

The definition which I give earlier as a mathematical definition is more applicable for theoretical problems, where we can see that the conditions of the are satisfied. Now, in this also there may be some discrepancy; for example, we are taking the ratio, now that ratio will have a limit, now the limit could be 0 or 1 also.

(Refer Slide Time: 51:43)



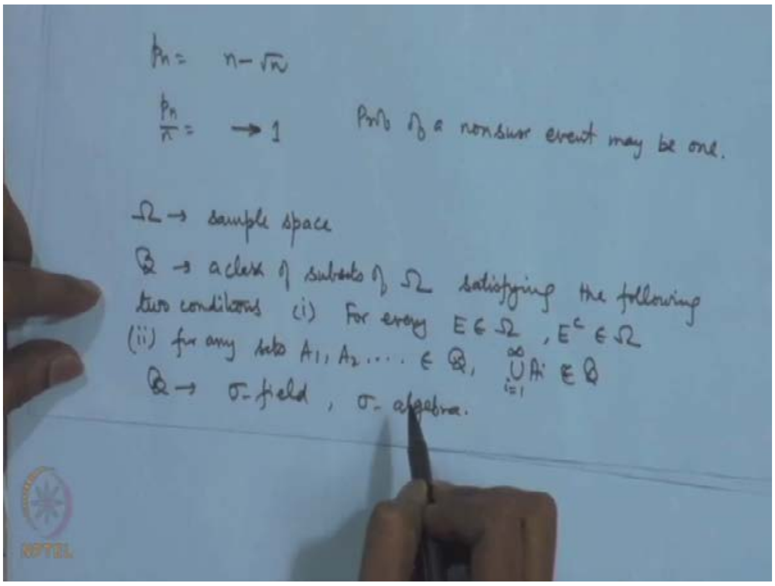
For example, you may say p n is equal to say n to the power1 by 3; now n to the power1 by 3 is not a negligible number, but if I consider p n by n then that will become1 by n to the power 2 by 3. So, that will goes to 0. So, probability of a non null event may be 0; which looks little counter intuitive. Although in the long run it has a meaning, what it mean that as n becomes large, n to the power 1 by 3 becomes much smaller compared to n. Similarly, we may have say p n is equal to n minus root n, in that case p n by n that will converge to 1.

Now, here you can see that the event is not a sure event. So, probability of a non sure event may be one, which is again little counter intuitive. Although from the is statistical distribution of the probability this is alright, because what did says that if we have n large, then out of that the number of occurrence of the event is almost full. That means, sometimes it may not occur, but that number is negligible; however, because of these drawbacks, we cannot use these definitions as the you can say arithmetic definitions, because they do not satisfy all the conditions. There are certain other problems also for example, this requires that the experiment be performed or we should be able to observe the occurrences, and the outcomes of the occurrences.

Now, they can be various experiment, where this is not possible. For example, rare phenomena or suppose we are looking at industrial experiment; and an industrial experiment suppose it is a large scale industrial experiment, in that case if we have to look at whether the system will fail at what time it will fail. Then suddenly, we are wait till the time in the system actually fails. So, in many of these conditions, the direct application of the definition is not possible. Based on this the (( )) Russian mathematician gave a the now the well known arithmetic definition of probability.

So, we have the sample space, and we consider a class of subsets of a class of subsets of omega; that means, these are events. Now, this should satisfy satisfying the following 2 conditions. One that for every E belonging to omega E complement also belongs to omega; that means, it is close under the complementation. And second is that for any sets A1, A 2, and so on belonging to B union of A i also belongs to B. That means, it is under closed closed under the infinite union also.

(Refer Slide Time: 51:43)



Such a B is actually called, this is called a sigma field or sigma algebra. This is a algebra is a structure; however, we are not getting into this. The main purpose is that, when we are considering a random experiment and its outcomes, then all the events should be included in the subject under study. That means, whatever set we are considering it should include all the (( )) events, and that is why we make it closed under the operation of complementation, and the infinite union. As we know that, this will further allow infinite intersections, it will allow the differences; that means, all the possible said theoretical manipulations will be included in the relevant is space.

And therefore, and therefore, we will be able to contact the study of the probability; that means, we can find out the probabilities of the related events provided, a certain probabilities are known to us in advance either by the first definition, that is the classical definition or by the second relative frequency definition. That means, probability of certain events may be known to us, and there after we can use them today the probabilities of various other events. So, in the following class, I will be discussing the axiomatic definition of the probability its ramification, and then various important results of probability.