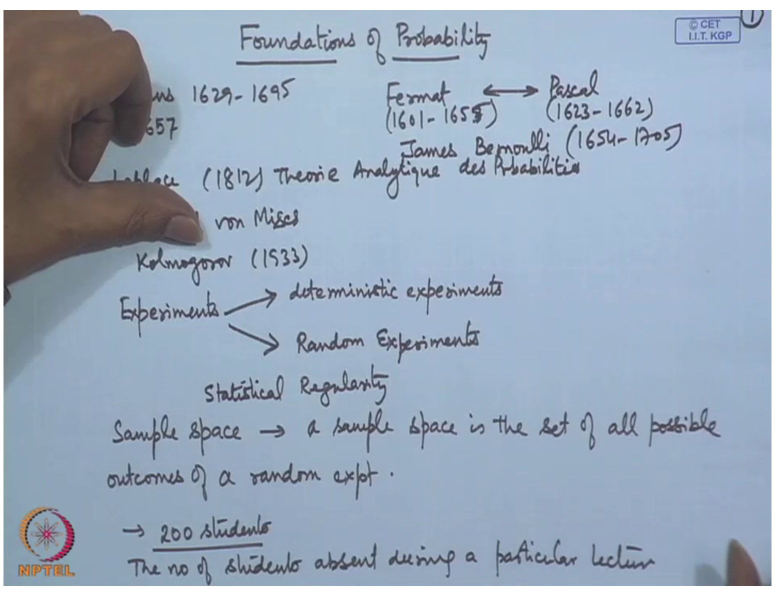
Statistical Methods for Scientists and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology - Kharagpur

Lecture - 01
Foundations of Probability

In this course Statistical Methods for Scientists and Engineers, I plan to cover several topics which are of one of the you can say most important topics for practicing scientists and engineers. So let me just say a few of these terms here, we will be introducing the term Probability and we talk about random variables and probability distributions. Then we discuss the main statistical methods that is of parametric point estimation and testing of hypothesis.

We will introduce one of the most widely applicable methodologies that is known as fitting of linear models under the term regression analysis, in planning of the experiments we considered experimental designs and we do the multivariate data analysis, and lastly we will cover nonparametric methods. So this is the outline of the course which we will be available on the website also.

(Refer Slide Time: 01:42)



Let me introduce the term probability, so we start with the foundations of probability. Loosely speaking we can say that the term probability is probably as old as the civilization itself, because people have been talking in terms like it is very likely that it will rain today, the winter seems to be colder than the last winter, and it is likely to go progressively more cold in the next year. It looks as if all the chances are that the food production this year will be more than the last year.

It is likely or there is a chance that the patient who is being operated will survive the operation, so this type of sentences or terminology have been in use. However, and suddenly it leads to the understanding that there is something called chance, probable, likely. However, a scientific study of probability theory started around 16th and 17th century in Europe during the initial period and probably the first published work in this direction is by Huygens, whose duration was 1629 to 1695.

And his published work in 1657 described the term Probability and gave a possible explanations of various type of events for which the probabilities can be calculated. Let me pint out that the formal study of probability started with the discussion of the gambling games, so that time in the Europe the people who were involved in the betting, gambling, coin tossing, die throwing, casinos etc.

They approach the mathematicians of that day to answer questions about that on which event they can bet so that their winnings will be more, and thus started famous correspondence with the 2 of the famous mathematician of that day that is Fermat and Pascal, the timelines of Fermat is 1601 to 1655, and Pascal live from 1623 to 1662, in the correspondence between these 2 mathematicians the first formal discussions of probability theory where there.

And it slowly took route especially the work by one of the Bernoulli that is James Bernoulli, you can say that the first formal definition of the probability came in Laplace famous work in 1812 that is Theory Analytic des Probabilities, later on his definitions and approach was considered to be insufficient to answer many questions. And therefore, the other approaches were introduced especially the empirical approach by Richard von Mises.

And finally the mathematical for mathematical definition that is by A. N. Kolmogorov Russian mathematician was published in 1933. I will be briefly discussing these things in detail, so let us consider that what is the basic terminology for Probability and what makes us study the subject probability. So in scientific theory we deal with experiments, so we have experiments which we can classify into 2 parts for the purpose of this discussion into what is known as deterministic experiments.

What are deterministic experiments? For example, I am holding this pen here and I put my pen on the paper, then the outcome is that there is a ink flowing through it and it makes certain impression on the page, this is a deterministic experiment. If we switch on a bulb then the bulb lights up, if we open our mouth to speak then words come out and others can listen.

If we consider experiments which are done in the classroom situations or in the laboratories situations many of the chemical experiments, you mix 2 chemicals certain molecules of one chemical with the certain molecules of another chemical you get another molecule. You consider some genetic experiments you have genes of certain thing and you mix up with genes of another one, you can use a gene for transplanting to treat a disease.

So these are the experiments which are of deterministic nature that means if we fix the conditions under which the experiment is performed the outcome is also fixed in advance that means it is known that what could be the outcome. For example, if we put a vessel full of water on a gas stove and we light the gas stove, when the temperature of the water reaches 100 degree Celsius under a certain atmospheric pressure the outcome is that the water will boil.

And similarly, the boiling temperatures for various other commodities for example milk, tea or any other liquid is there, for example you may even melt iron if you put it in a furnace with a very high temperature something like 1600 degree Celsius, these are examples of the deterministic experiments. However, in probability theory we are concerned with the experiments which are termed as Random experiments.

Now what is it that separate random experiments from the deterministic experiments, as I mentioned the deterministic experiments, if we fix the conditions under which the experiment is being conducted then the outcome is known in advance or it can be fixed in advance? However, there are other situations where even if we perform the experiment under fixed conditions, there are various other uncontrolled factors which make it impossible to predict the outcome of the experiment.

So for example we study weather, so weather is a subject which is study by physicists, meteorologist and so on, even then every year what would be the average rainfall during the season? What would be the total rainfall during the season? What would be the total what would be the average temperature during a month? All of these things are varying that means we cannot predict with certainty that this is will be the thing.

For example, we cannot say that during the monsoon season India, Indian subcontinent or a particular town in the continent will get say 20 centimeter of the rain, we cannot say with certainty, it maybe it is 15 centimeter, it maybe it is 30 centimeter and so on. A similar thing is about temperatures, a similar thing is true about say agricultural product we may fix up a size of a farmland on which the seed of a certain crop will be sown.

We may fix what kind of pesticides will be used, what type of fertilizer will be used, what with what frequency irrigation facilities will be provided and so on, even then we cannot say that the total food grain production say wheat or rice or maze from that particular farmland would be say 100 metric ton or 50 metric ton or 10 metric and etc. we cannot exactly say that how much would be the food grain production it will be in certain range.

We have a bulb, so when we switch on the light I mean the lights switch, we have the outcome that bulb will be lighting, however, what will be the total lifetime of the bulb that cannot be predicted in advance, it may be 10 hours, it may be 10 days, it maybe even a year, a bulb life can be varying as much. So these are the examples where even if we perform the experiments under fixed conditions.

For example, the bulb is produced mechanically using machines with a certain material the tungsten and so on, all those things are fixed even the place where we are using the light bulb the wiring the light switch everything can be controlled, but even then how long the bulb will light, will keep on giving the light is not known in advance. So we observe that a large number of natural phenomena in science in engineering or having the outcome which is unknown.

And that makes it interesting to study the subject probability, because he may feel that if the science is done on with certain hypothesis and certain conditions then the results are fixed, however, that is not so, even something like making of a road or making of a bridge on the river, so that is a strictly an engineering event. For example, you construct the bridge using certain material and make it using certain standardization.

But even then the total life of the bridge is it 100 years, or is it 120 years, or is it 20 years, we cannot say in advance because that will be dependent upon the weather conditions, the number of floods that bridge may have to face, and the amount of traffic which may vary too much I mean it may be that the traffic is much less in certain years, and then and suddenly it goes up, all of these things will determine the life of the bridge.

So that is how are we can say that the reason for studying the probability theory is that most of the natural phenomena which look like that one could have formed from scientific rules for that, but they are actually uncertain in nature. And therefore, we need to look at the possibilities of various outcomes, and then we need firm foundations for doing so that means what are the rules by which we can calculate the probabilities of various outcomes.

So it is not that only theoretical experiments like tossing of a coin, throwing a die, or picking of a card from a pack of cards is the example of random experiments, almost every happening in the natural phenomena is actually example of a random experiment. For example, birth of child what would be the life of a person that means the age of a person, so a child is born but what would be his total age, it can be 5 years, it could be 10 years, it could be 60 years, it could be 90 years and so on.

So almost all the activity in human you can say under which we can consider, we can conceive of they are actually part of the random experiment. And therefore, we need to study the subject probability. Another thing which is noticeable here is that when we talk about say phenomena for example rainfall, so we may not be able to say exactly whether the rainfall will be 100 centimeters or 120 centimeters etc.

But over a long period of time if we have observed this for several years, then we may be able to say that what is the probability that the rainfall will be <100 centimeter or >say 120 centimeter or between 100 to 120 centimeters that means when we are studying any event in a probabilistic way we need to look at the long-term behaviour of the event, it is not one of that suddenly I asked the question that what is the probability that I would collapse while taking this lecture?

So this could be a one of the event, because it has not happened and it has not been observed, so one may not be able to tell the probability of this event, what is the probability that suddenly this roof will collapse? Where I am taking the lecture, so one may not be able to answer those questions in a satisfactory way. However, how much time I would take to complete the topic that I am teaching today, I can say with almost certainty that I will take 2 lectures or maybe I can say that it is around 120 minutes, so I can fix up a range 110 minutes to 130 minutes.

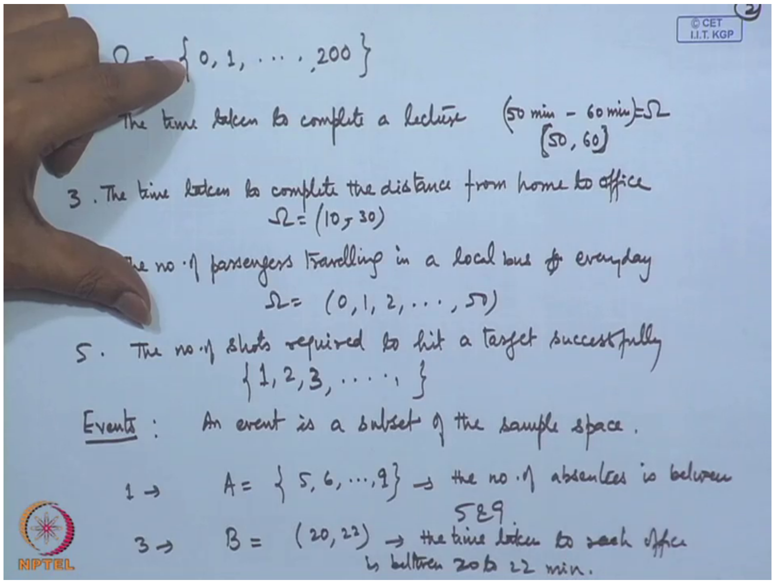
In that case this is my outcome of taking lectures on Probability over several years for students at IIT Kharagpur where I am actually teaching, so I know that this topic I usually finish in 2 lectures, so I can say with almost certainty that in 2 lectures or you can say with the little variation that 2 lectures are equal to 120 minutes, so between 110 to 130 minutes I can complete this topic.

So this long-term behaviour which is called statistical regularity allows us to study the subject probability, because over a long term when we study the behaviour of the events, one may be able to say what would be the probability or what is the likelihood of a certain event happening. Now I will start with the formal terminology of the random experiment on which I can define the probability. So for that I start with what is known as sample space.

So a sample space is the set of all possible outcomes of a random experiment, let me look at a few examples here. My lecture on probability it has say 200 students okay, but everyday all the

200 students do not turn up, in a semester I take around 44 or 45 lectures, in each lectures few students are absent, so I consider the experiment as taking of a lectures and whether the students how many students are coming.

(Refer Slide Time: 19:07)



So if I look at the outcome that how many students, the number of students absent during a particular lecture, if this is my random experiment then the sample space can be written as so theoretically speaking I can say nobody may be absent, 1 student may be absent, and if I am saying that my class consists of 200 students then on a rare occasion it may be happen that all 200 are absent okay.

That event may correspond to some lack of communication like I may feel that it is a working day but actually it may be declared holiday which I may not know and therefore all the students may not turn up for the lecture, so the number of possibilities here I can list as 0, 1 up to 200. Let us consider another example say the time taken to complete a lecture, now 55 minutes are allocated for taking a lecture.

However, when I am coming towards the end of the class I might have started a topic which I need a few more sentences to speak before I end the lecture or I may be able to complete it just before 55 minutes and therefore, after that when there is a new topic I may not need to introduce it at that stage. So I may take say 50 minutes to 60 minutes just to keep the upper bound because after 60 minutes I cannot continue because by that time the students of another class would have entered.

So I may consider my sample space to be 50 minutes to 60 minutes' interval that means I may write it as open interval or closed interval 50 to 60. Let us consider the time taken to complete the distance from home to office, now there may be the total distance maybe fixed, for example it may be 5 kilometers, but there may be traffic on the road, there may be railway crossing on the way, and therefore, the time taken to complete this distance may vary quite a lot.

For example, a 5 kilometers distance one may cover in 10 minutes, but on other occasions one may take 12 minutes, 15 minutes depending upon the traffic or as I mentioned there may be a railway crossing on the way which may be closed sometimes, so we may safely put say 10 minutes to 30 minutes time may be taken to complete the distance from home to the office, so my sample space here in this case can be an interval 10 to 30.

The number of passengers travelling in a local bus every day, so there is office bus for example so once again it may have the total capacity say 50 and the number of people who may be travelling in this bus maybe from 0 to 50 on each day. We are targeting, there is a target practice and the number of shots required to fire or to hit a target successfully, so one may be successful in the first shot, he may need second shot, he may need third shot and so on.

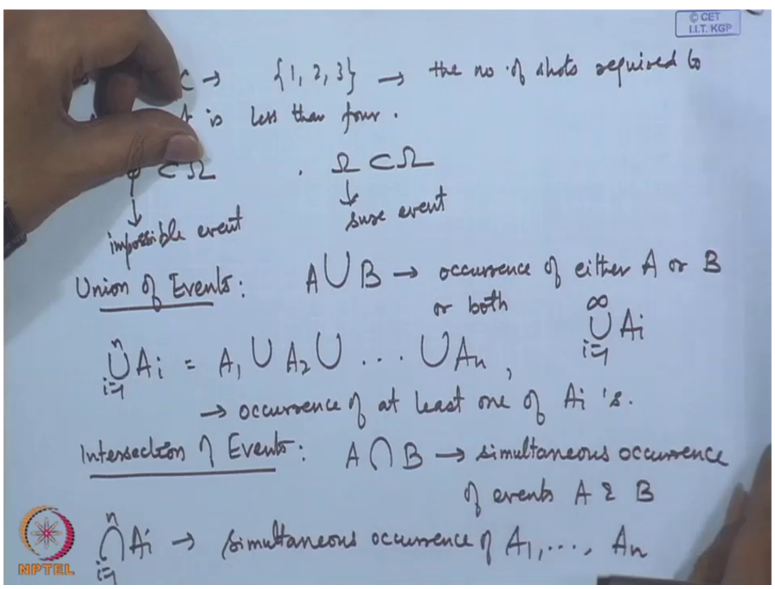
So I am just putting and so on to indicate the possibility that one may never be successful. There are certain things for example if we consider finding out a sure shot treatment of a disease such as say cancer, then their trails have been going on since time immemorial, but we have not been able to come up with readymade solution or you can say fixed solution which will work for all the instances of this disease, although there are cases where that disease is cured.

But then there are many other cases where the disease is not cured even if the same treatment is given, so there is no sure shot you can say solution or treatment which will treat all the instances of this disease, so if we are looking for a ultimate solution for this problem, then may be the number of trials is infinite as far as we are concerned today. I have given examples of various natures here; you may look at here the number of entries in the sample spaces is finite.

This is infinite it is an interval here, this is countably infinite, so that once again tells that the ways of describing the sample space can be many and the number of entries in the sample space can be finite countably infinite or it could be uncountably infinite. So we talk about events, so an event is a subset of the sample space, so for example if I say in the first experiment I consider A as 5, 6 up to 9 say, what does this denote?

This denotes the event that the number of absentees is between 5 and 9. I may in the say let us consider third experiment the time taken to complete the distance from home to the office, I may consider the event as say 20 to 22, so this means that the time taken to reach office is between 20 to 22 minutes.

(Refer Slide Time: 26:29)



See in the experiment of looking at the number of shots required to hit the target, suppose I define the event say C by saying 1, 2, 3 that means the number of shots required to hit a target is <=3 or you can say <4, so these are various events. So when we consider the subject probability we are interested in the probabilities of various events, now one may just asked the question that whether we should consider all types of events?

Then certainly one answer is that if we consider all subsets of the sample space then that will consist of all the events that means we can consider the power set of the sample space that would be the ultimate event space, but I will show you later that it is not necessary that we consider all the events in each instances of a random experiment, because one may not be interested to enlist everything. However, let us discuss various kinds of subsets.

So for example phi is a subset, so just to say that we can use certain notations, for example I have used notation omega for a sample space, so varies notations for the sample space sometimes people use capital S, sometimes people use theta and so on, the various notations are used for the or sometimes the universal set is denoted by U etc. so any subset of that is denoted usually by the capital English letters like A, B, C, D and so on, these are the useful terminology for the events.

So naturally empty set is a subset of every set, the full set itself is a subset of itself, so these also correspond to certain events, this corresponds to impossible event, this corresponds to sure event that means this is certain to happen and this will never happen. Now we when we have interpreted the events in terms of the sets then we can now use the framework of the set theory, for example when we talk about the sets we talk about certain algebraic operations on the sets.

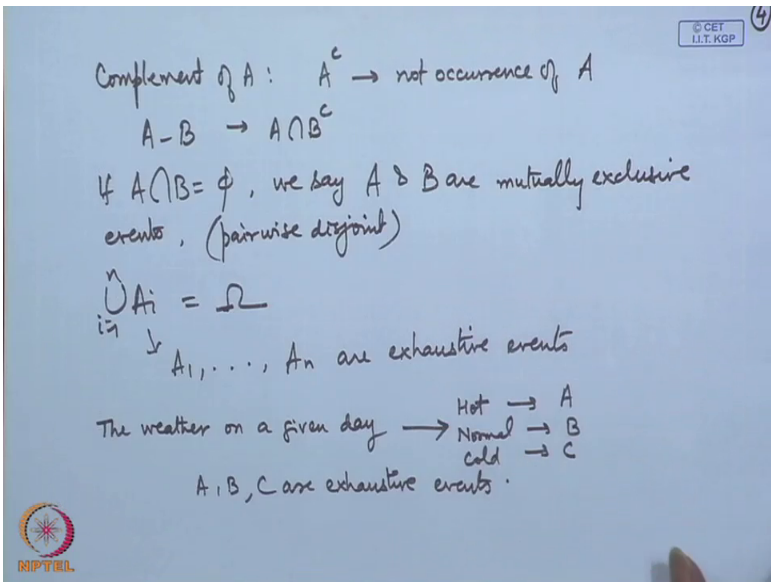
For example, Union, Intersection, Complementation, taking difference and so on, so these will also correspond to certain events let me explain this. So for example union of events, so if I have 2 sets A and B, then A union B is the collection of the elements which are either in A or in B or both. Now when I say A and B are events that what does A union B will denote? It will denote that occurrence of either A or B or both.

Now likewise I may consider more than 2 events, suppose I have 3 events A, B, C then A union B union C makes sense, because it would mean that happening of either of A, B or C are either of 2 of them or all the 3 of them. So in general we can talk about union Ai i=1 to n, which we write as A1 union A2 union An etc. this is occurrence of at least one of Ai's at least one of them occurred like 1 may occur 2 may occur and so on all the n may occur.

Similarly, we can talk about the intersection of events, in set theory we know that A intersection B is a set of those points which are common to both, here it will mean simultaneous occurrence of events A and B, likewise one can talk about intersection of Ai i=1 to n, this will mean simultaneous occurrence of A1, A2, An. One may talk about in finite unions and in finite intersections also.

For example, one may talk about union Ai i=1 to infinity, this would mean occurrence of at least one of Ai's. And in a similar way one may talk about intersection of Ai i= 1 to infinity, so this would mean simultaneous occurrence of all the events A1 A2 and so on.

(Refer Slide Time: 31:39)



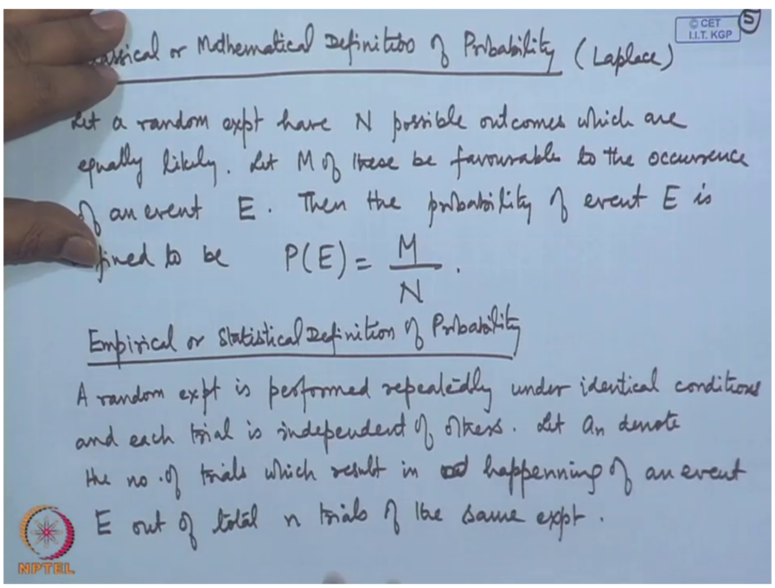
Similarly, we have the complementation for example complement of A so that is A complement, now in set theory complement is the collection of those elements which are not in that set but they are in the universal set. So here it would mean not concurrence of A that means A has not occurred so the complement of A takes place, now using this one can talk about anything else for example if I say A-B then this is having A intersection B complement that means happening of A and not happening of B and so on.

So one can basically now consider the interpretation of all types of set operations in terms of events. Now we have special cases for example we have disjoint sets, now if I say disjoint sets they are corresponding to events what does it mean? It means that there is no element common, if that happens then statistical speaking it would mean that if the event A occurs, then the event B cannot occur or if the event B occurs then the event A cannot occur.

So this is known as mutual exclusion, so we call such things as so if A intersection B=5, we say A and B are mutually exclusive events that is we can also use the terminology pairwise disjoint that means the two terms things taken together are disjoint we also have another thing for example a certain number of sets the union of them may be =omega, if that is happening then we say A1, A2, An etc. are exhaustive events.

Because it means that all the possibilities of the sample space are taken care by A1 A2 and An, for example if we are considering say for example the weather on a given day, and I have descriptions like it may be dependent upon the weather, so that means that type of conditions that we may have, we may have hot, we may have normal, or we may have cold. So we can use denotations here A, B, C. So then this will mean that A, B, C these are exhaustive events.

(Refer Slide Time: 35:19)



Let us go back to the development of the subject as I mentioned, and one of the first you can say formal introduction or formal developments of the subject was published by the French mathematician Laplace in 1812 in his book Theories des Analytic Probability, and one of the first formal definitions was given by him which we now term as classical or mathematical definition of probability, so this is by Laplace.

Let a random experiment have say N possible outcomes which are equally likely. Let M of these be favourable to the occurrence of an event E. Then the probability of event E is defined be M/N. So this is one of the earliest definitions and it is applicable to experiment such as coin tossing, die throwing that is experiments which actually originated the mathematical treatment of the subject probability.

So in those cases one may assume that this type of conditions will be satisfied, like we toss a coin so we will get an equal probability half, half for a head and tail. Similarly, if we have a die and we roll the die, then we allocate probability 1/6 to each of the impossible spaces 1, 2, 3, 4, 5, 6 coming up, so this is based on the assumption that the coin is fair, the die is fair and so on. This is also the drawback of this definition because suppose the coin is not fair then this definition is not applicable.

The definition is also not applicable to the cases where we cannot describe the all the outcomes in a proper way, the number of outcomes may be infinite, just before this I was talking about several random experiments we considered for example the number of shots required to hit a target successfully here the number of possibilities are infinite countably infinite. If you look at the time taken to complete the lecture so it is an interval 50 to 60 it is uncountably infinite.

In all of these cases this type of definition is not applicable, so the definition has limitations. So therefore, it was said that or it was felt that probabilities definition should be based on empirical evidence, as I was mentioning that we may not be able to say that whether the rainfall this year or next year would be <100 centimeters or not, but based on single thing but have observed the weather over past 50 years or 100 years.

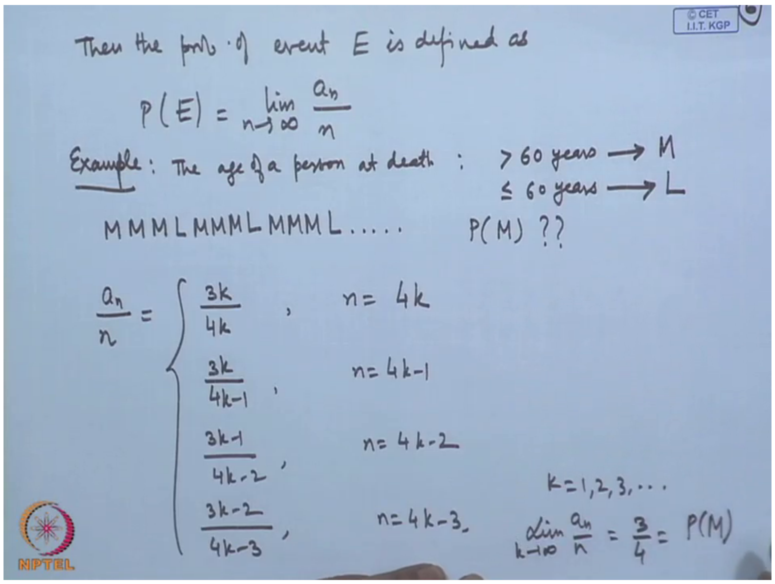
Then we may say that out of previous 100 years say 60 time it happened that the weather was rainfall was >600 centimeters, and over say 80, we are able to say something over say 70, we are able to say something over 50 years we were able to say something, therefore, we can fix this number 60/100 that is 6/10 or 0.6 as the probability that the rainfall will be >100 centimeters in the coming here. Now this is based on the evidence so this is called evidence-based definition or empirical definition of probability.

So I will just give this thing here now empirical or we can say also statistical definition of probability, if a random experiment is performed repeatedly under identical conditions and each trial is independent of others, that means the outcome of a particular trial of a random experiment is not affecting any other incidents of the same experiment, that means if the experiment is performed several times then what happened in one of the trials does not affect what happens in the other trials, so this is called independent.

And under identical conditions means that the experiment is exactly the same, for example if I consider a tossing of a coin, tossing of a die, or if we are considering taking off the lectures, so we consider the conditions to be identical. So if that is so the experiment is performed repeatedly under identical conditions and each trial is independent of the others now let us consider some 10 numbers let a n note the number of trials which result in happening of an event E out of total n trials of the same experiment.

That means the experiment is performed n times of out of that a n is the number of trials in which we can say that the event E has occurred.

(Refer Slide Time: 42:07)



Then we define the probability of event E is defined as limit of a n/n as n becomes n tends to infinity, it means that if the ratio of the number of happenings which are favorable to the event E to the total number of occurrences of the random experiment, if this limit exists then this limit is assigned the probability of the event E. Let me just give an example to show that how you can actually use this in reality.

I just mentioned that weather experiment that whether the probability that the rainfall will be >100 centimeter, so what is the probability that the next child will be born will be a boy or a girl so you usually associate probability half, so that is because the past experiences says that usually the child birth when they are happening they are equally likely to produce the boy or a girl. Similarly, if we say that the average life of a person is in say USA is 72 years so of a male for example.

So it is based on the data of the mortality over a period of time and it may be the current thing, because the average longevity has been increasing earlier it was say 60 years, and it will become 62 years and so on. So let us consider an example here, so I am considering the age of a person at death okay so it could be say >60 years or <=60 years, if it is >60 years I call it an event M, if it is <60 years I call it an event L.

And I observe the sequence of deaths at a particular town, and it is like the data comes from hospital or from mortuary etc. so suppose the sequence results in MMM this is artificial sequence here L that means the first three persons I observed they died at the age >60 years, then the next person <60 and so on, and so this artificial sequence for example it results in 3 being >60, one <60 and following the sequence like this.

I want to calculate what is the probability of M, of course by looking at this sequence you should say that it is 3/4, let us see that our empirical definition yields the same or not, so we can look at this what is the ratio of the number of happenings favorable to the event M to the total number of occurrences of the number of trials here. So here you look at this I can express as something like 3k/4k if n is of the form 4k, it is 3k/4k-1 if n is of the form 4k-1, it is = 3 k-1/4k-2 if n is of the form 4k-2, it is of the form 3k-2/4k-3 if n is of the form 4k-3, for k=1, 2, 3 and so on.

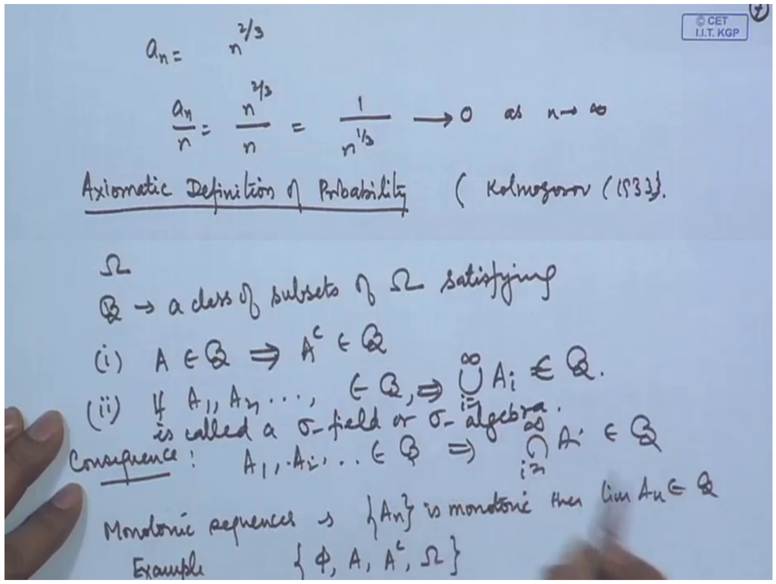
So if you look at this then limit of this sequence here, if you look at this, this is 3/4, this is 3/4, this is 3/4, this is 3/4, as k tends to infinity, so we can say that the probability of a person achieving age >60 is 3/4. So here you can see that it is empirical definition being applied here, so likewise for most of the real life phenomena which I was discussing in the beginning of this lecture, in all those situations one can apply empirical definition of probability to obtain probabilities of various events.

So one may say that this definition is almost you can say universally applicable, but even then it has certain limitations. For example, one should be able to observe the experiment although you may not be able to conduct the experiment, but you should be able to observe the experiment and what is the outcome of that, and sometimes that may yield falsified thing, for example you may observe the outcomes but you may not be able to look at the outcomes in its entirety.

And therefore, some data for example you have put some machine to record and that machine is malfunctioning, or some person a human being is collecting the data and he may get a wrong figure, so these are the drawbacks of this empirical event. Secondly, one may not be able to conduct the experiment, as I was mentioning that the experiments are rare which may not happen so often, then also the application of this definition is not possible.

For example, I was mentioning one of events, so in those cases if somebody has not observed that kind of events then one cannot find the probability of that. Another thing which could be misleading is for example one may feel that if it is an impossible event the probability is 0, the converse should also be true, like if the probability is 0 the event should be impossible. And similarly, if the event is sure you have probability 1, but if the probability is 1 the event should be sure event.

(Refer Slide Time: 48:30)



But one we have this kind of occurrences, for example I may have a n=n to the power 2/3, and so if I look at the ratio a n/n then it is n to the power 2/3 added by n that is = 1/n to the power 1/3, now this goes to 0 as n tends to infinity, so this would mean that the probability for this event for which the number of favourable outcomes is n to the power 2/3 is actually is 0, but this is not an impossible event actually the event occurs but the ratio of the occurrences this converges to 0.

So one may give interpretation in such a way that the number of occurrences is actually progressively decreasing like if I say n=1 then a n is also 1, but if I take n=2, n=3, n=8 for example then a n=4, if I take n=27 then this will become 9 that means the ratio is becoming much less and progressively it is declining to 0. So it is not an impossible event but in the long run the probability of occurrence of the event would be negligible so that is the meaning of this.

And only have the reverse of this also that the probability maybe 1 but the event may not be a sure event. Now to overcome these drawbacks of both the definitions although we may have considered them as now the methods of calculation of the probability, the Russian mathematician A. N. Kolmogorov laid the foundations by giving his axiomatic definitions, so for this one now we have a sample space so omega is the sample space.

And as I mentioned that by events we mean that any subset and one may consider all subsets also, but in a complex random experiment it may be quite complicated or it could be quite difficult to enumerate all the events and then look at the possibilities of that, and it may not be of much interest also. For example, when we are discussing the rainfall, the amount of rainfall then certainly it could be like it could be a drought that means complete drought is there.

It could be 10 centimeters, it could be 15 centimeters, it could be 200 centimeters, there can be a super cyclone in that period and so on, there can be thunderstorms lot of possibilities are there and if one wants to study, it could be a very complex description of the sample space, but an average person, or an average weather man, or an average farmer may not be interested in all of that thing, he may be interested only in the information whether there will be an adequate rainfall or not.

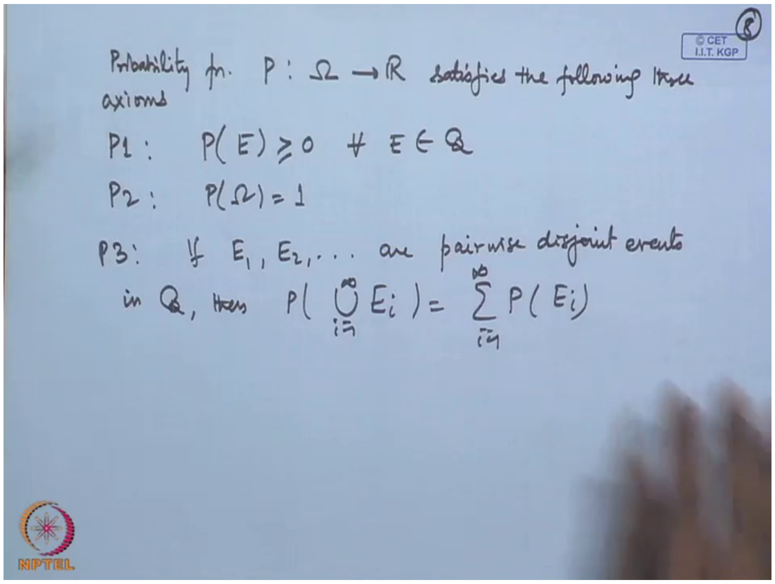
So if we say that then I am looking at only 2events I may say A is the event that the rainfall is adequate, A complement is the event that the rainfall is not adequate, and I am not bothered about anything else. Therefore, we should have a framework in which we can limit the number of events that we may be considering, so we consider a structure B so I use a notation script B, it is a class of subsets of omega satisfying the following 2 assumptions.

One is that A belongs to script B implies A complement belongs to B that means for every event it is complements should be there. Secondly, if A1, A2 and so on belongs to B then its union must also be in B that means for any collection of the events its union will also be there. As a consequence, one may check that if I say A1, A2 and so on belongs to B, then this mean that the intersection belongs to B.

We may consider monotonic sequences, then if An is monotonic then limit of An also belongs to B and so on, so all types of possibilities are there. So simplest example is I just now mentioned I may consider phi, A, A complement and omega, then this satisfies this. For example, A union complement is omega, A intersection A complement is phi, A union phi is A and so on, all possibilities will be there.

This type of structure this is called this structure is called a sigma field or sigma algebra. So what I am saying is that we need not consider the class of all subsets of a sample space, but we may restrict attention to few events which are of use to us, and they should formulate a or they should form a structure that structure we are calling a sigma field or a sigma algebra.

(Refer Slide Time: 54:31)



Then we define probability function P is from omega into R satisfies the following 3 axioms, let me call it we P1 that is the probability of E is >=0 for all E, P2 probability of omega is 1, P3 is that if E1, E2 and so on are pairwise disjoint events in B, then probability of union Ei i=1 to infinity= sigma probability of Ei i= 1 to infinity. So these are called axioms of probability and now I am just defining P to be a function which just satisfied this okay.

Now we will show that this structure is enough to consider the probabilities of various kind of events, we will see the consequence of this in formulation of certain rules of calculation of probabilities of various events, for example probabilities of unions, probabilities of intersections, we will use it to define conditional probabilities and various other things. So in my next lecture I will elaborate on this axiomatic definition how to use it and we will discuss certain examples, before moving to the concept of random variables and probability distributions.