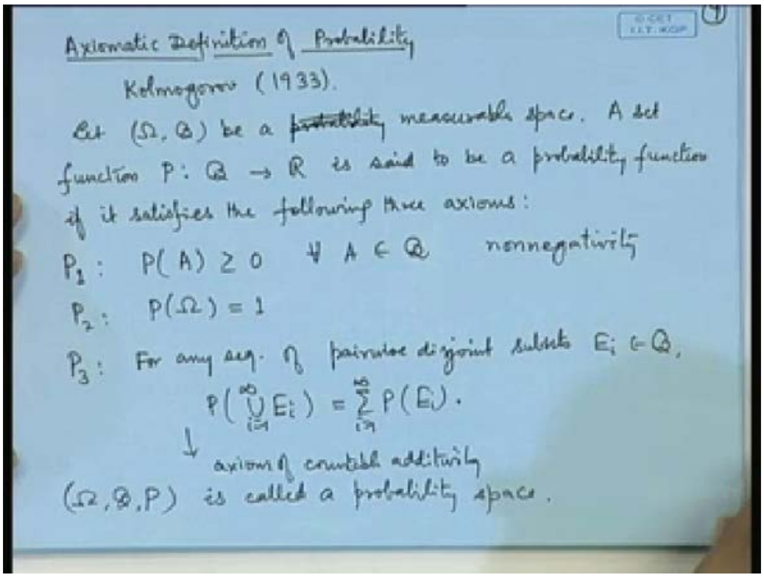
Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur
Module No. #01

Lecture No. #04
Laws of Probability-I

In the last lecture, I have introduced the axiomatic definition of probability.This takes care of thedeficiencies or drawbacks left by the classical definition or the relative frequency definition of probability. So, in this definition we give a general framework under which a probability function is defined. Thisdoes not tell you how to calculate a probability. But, a probability function must satisfy these axioms in order to be a proper probability function.

(Refer Slide Time: 00:23)

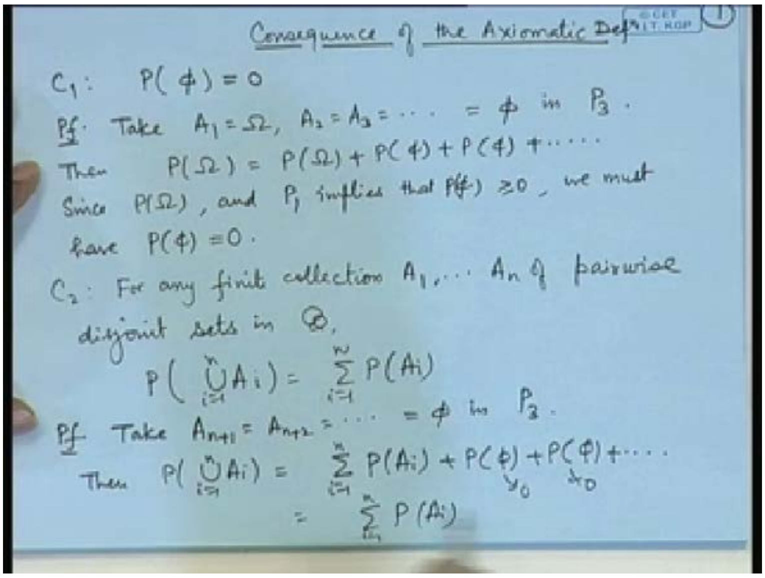


So, in particular if we have a sample space and a sigma field of subsets of that sample space let us call it as script B then a set function p from b to R is said to be a probability function if it satisfies the given 3 axioms which we name P1, P2, P3. Thefirst is the non negativity axiom; that is, the probability of a is greater than or equal to 0 for all A belonging to B.

So, this is the axiom of non negativity; then probability of the full sample space is 1.

Basically,it makes the P2 be a finite function andthe third axiom is the axiom of countable additivity. Thatis, for a given pair wise disjoint sequence of sets probability of the union is equal to the sum of the individual probabilities. Thus, this omega B and P is called a probability space.

(Refer Slide Time: 01:59)

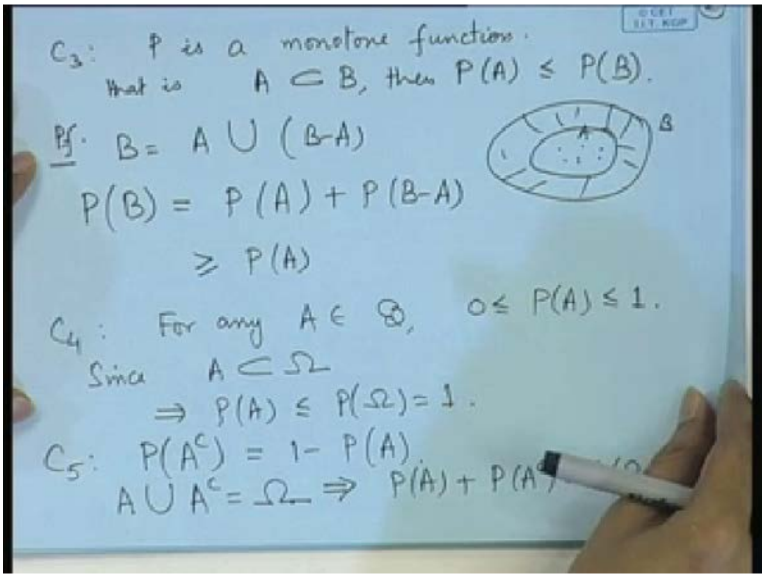


Now, some of the consequences of the axiomatic definitionare as follows: the first consequence is, let me call it, C1. Thatprobability of the impossible event must be 0. Toprove this statement let us take a 1 is equal to say omega and A2,A3, etcetera, to be phi in axiom P3; then, we will get probability of omega is equal to probability of omega plus probability of phi plus P of phi plus P of phi, etcetera. Sincep omega is 1 and P1 impliesthat P phi is greater than or equal to 0, we must have P phi equal to 0.

The second consequence is that for any finite collection A 1,A 2,A n of pair wise disjoint sets in B probability of union A I i is equal to 1to n is equal to sigma probability of A I i is equal to 1to n. Let me explain this that why do we need this finite additivity consequence here to be provedwe have assumed the countable additivity axiom, but that does not necessarily imply the finite additivity.

A proof of this can be given using the fact that in A3, we can take An plus 1A n plus 2, etcetera, to be phi in the third axiom. Thenwe will get probability of union A I i is equal to 1to n is equal to sigma probability of A I i is equal to 1 to n plus p phi plus p phi etcetera. Now, if you use consequence 1 here then, these terms are 0 and we get sigma probability ofA I i is equal to 1to n.

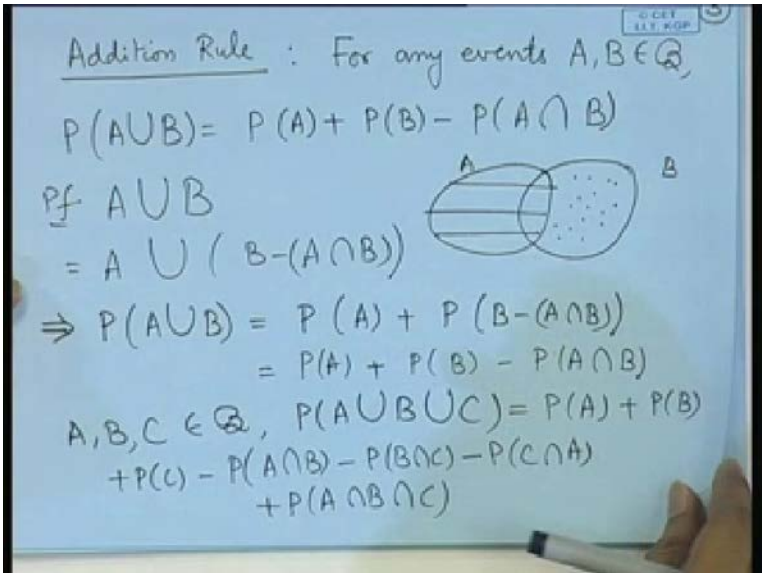
(Refer Slide Time:05:03)



A third consequence is the P is a monotone functionthat is if I take say A to be A subset of B then probability of A will be less than or equal to probability of B let us look at the proof of thisconsider say set A and A set b then I can write B as A union B minus A that is this is B minus A and this is A and these 2 are disjoint. So, if I make use of the finite additivity consequence then, probability of B is equal to probability of A plus probability of B minus A. Naturally, this is greater than or equal to probability of A since probability of B minus A is always greater than or equal to 0.

As a further consequence, we have that for any event a probability of A lies between 0 and 1. Now, the first part of this is always true because of the P1 axiom. Now,A is A subset of omega; for every omega for every A this implies that probability of A is less than or equal to probability of omega, that is equal to 1. If I consider probability of a complement then it is equal to 1 minus probability of a this follows because, I can write A union A complement as omegaand therefore, probability of A plus probability of A complement is probability of omega that is equal to 1.

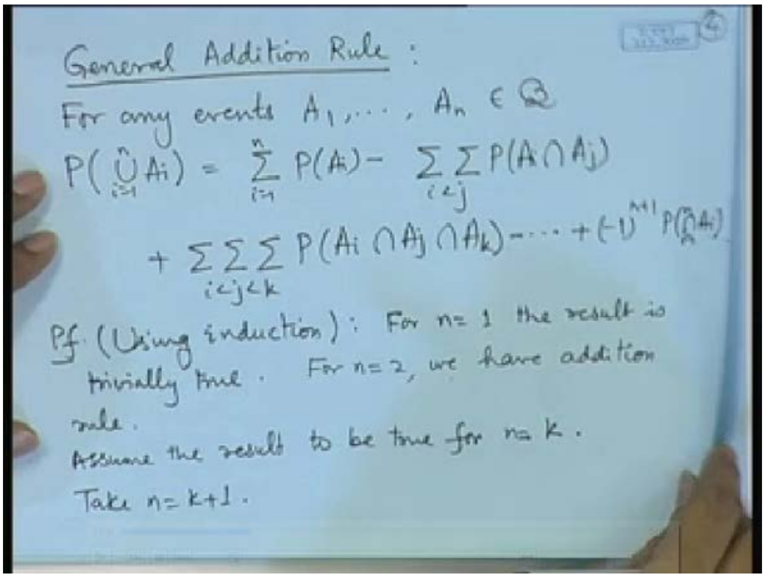
(Refer Slide Time: 07:25)



Now, we look at certain further consequences of the definition the first of them is the addition rule for any events A and B; probability of A union B is equal to probability of A plus probability of B minus probability of A intersection B in order to prove this statement. Letus consider any 2 sets A and B then A union B can be expressed ascA union B minus A intersection B. So, we can write A union B as A union B minus A intersection B; we can observe here that this is A disjoint union.Therefore, if I consider probability of A union B it is equal to probability of A plus probability of B minus A intersection B.Now, at this stage we notice that A intersection B is A subset of B and if we look at the statement that probability of B is equal to probability of A plus probability of B minus athen, this implies that probability of B minus A is equal to probability of B minus probability of A. That means, if A is A subset of B then probability of B minus A can be expressed as probability of B minus probability of A.Therefore, here we can write this as probability of B minus probability of A intersection B.

Now, naturally 1 can think of the generalization of this rule for example, if I consider say for 3 events suppose A B and C are 3 events then we must have probability of A union B union C is equal to probability of A plus probability of B plus probability of cminus probability of Aintersection B minus probability of B intersection C minus probability of C intersection A plus probability of A intersection B intersection C1 can look at this statement from the point of view of set theory or Venn diagram. If I consider 3 events say AB and C then the union can be expressed as A union b union C. However, here we have to remove A intersection B B intersection C and C intersection A.Ifwe remove that then the set A intersection B intersection C has been removed 3 times. So, we have to add it once to get this portion here. So, a intersection B intersection C has to be added here.

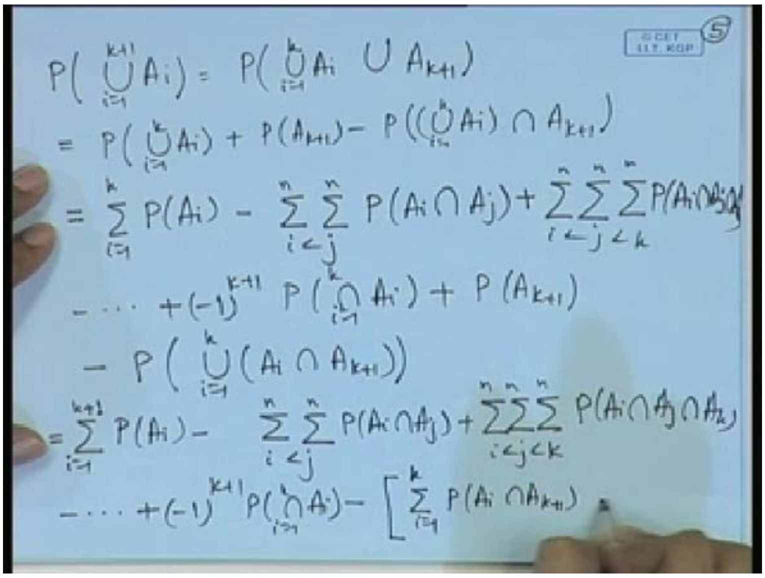
(Refer Slide Time: 11:13)



So, this gives us a rule for considering a general addition rule and we have the followingresult: general addition rule. So, if we have events A1A2A n then, probability of union A IIis equal to 1to n can be expressed as sigma probability of A I I is equal to 1to n minus double summation probability of AI intersection AJ I is less than J plus triple summationprobability of AI intersection A J intersection AKI less than J less than K minus and so on.Finally, you will have minus 1 to the power n plus 1 probability of intersection AII is equal to 1to n.

I can prove this result using induction; for example, if we take n is equal to 1 the result is trivially truefor extension from k to k plus 1. Wewill need the result for n is equal to 2 which has already been proved for n is equalto 2; we have addition rule. Now, assume that the result to be true for n is equal to k; now take n is equal to k plus 1. So, we need to consider probability of union AIIis equal to 1 to k plus 1 andwe can consider it as probability of union.A II is equal to 1 to k union A k plus 1.

(Refer Slide Time: 13:34)



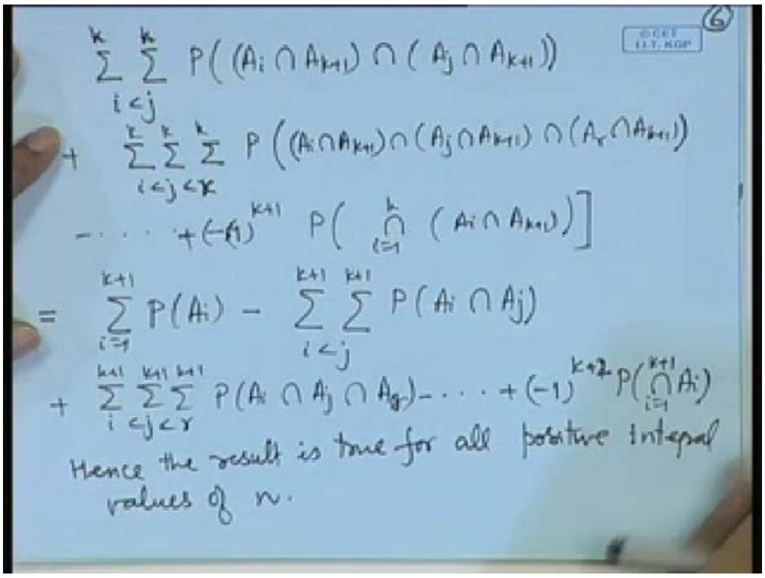
So, now I can apply the result for the union of A and B2 events. So, this becomes probability ofunion AII is equal to 1 to K plus probability of AK plus 1 minus probability of union AII is equal to 1 to K intersection A K plus 1. Now, the first part of this can be expanded because we have already assumed that this rule is true for n is equal to K. So, this becomes sigma probability of AI Iis equal to 1 to K minus double summation probability of AIintersectionAJI less than Jnow this sums are up to n triple summation probability of A I intersection a j intersection A KIless than Jless than K the sums are up to n and. So, on plus up to minus 1 to the power K plus 1 probability ofintersection A I*I* is equal to 1 to k.

Now, then we have probability of a k plus 1 and here we apply thedistributive property of the unions and intersections. So, this becomes minus probability of union A I intersection A K plus 1 I is equal to 1 to k now if you look at this last term it is again union of K events and since we have assumed theprobability of union result to be true for n is equal to k we can apply that formula.

So, using that we will get summation of probability AI intersection AK plus 1 for I is equal to 1 to k and that term can be adjusted with this. So, let me write it here. Firstly, sigma probability of A II is equal to 1 to k minusdouble summation I less than Jupto n probability of AI intersection AJ plus triple summation I less than J less than K probability of AI intersection A J intersection A K minus and. So, on plus minus 1 to the power k plus 1 probability of intersection A I Iis equal to 1 to K.

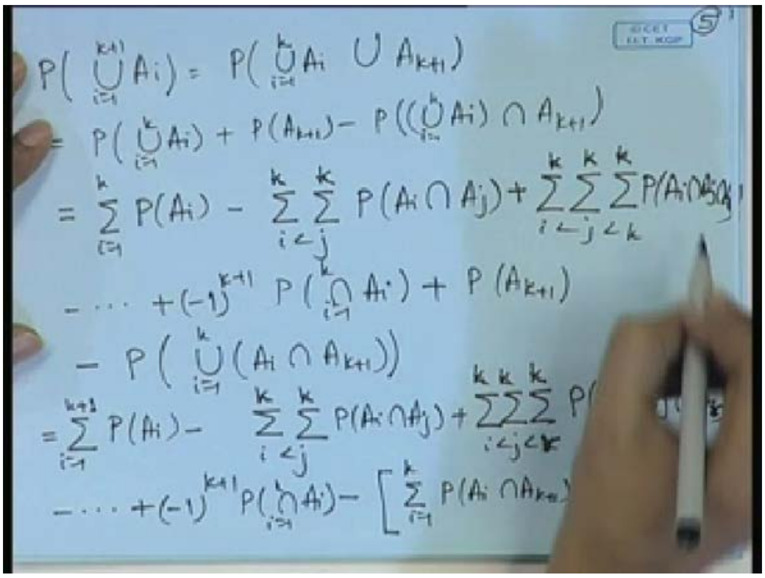
Now, this probability of AK plus 1 can be added to the first term. So, the first term becomes probability of A IIis equal to 1 to k plus 1 now let me expand the lastunionby using the formula for N is equal to K. So, this becomes sigma probability of AI intersection AK plus 1I is equal to 1 to K minus.

(Refer Slide Time: 17:06)



Now, you will have double summationprobability of AI intersection AK plus 1 intersection with AJA k plus 1 where I is less than J and the sum goes up to n only goes up to K. So, I thinkI have made some small mistakes here.

(Refer Slide Time: 17:31)



Thesesums are up to K, then you will have triple summation I less than J less than K probability of AI intersection AK plus 1. So, we may put it R intersection AJ intersection A K plus 1 intersection a r intersection AK plus 1 and so on minus 1 to the power K plus 1 probability of intersection AI intersection AK plus 1.

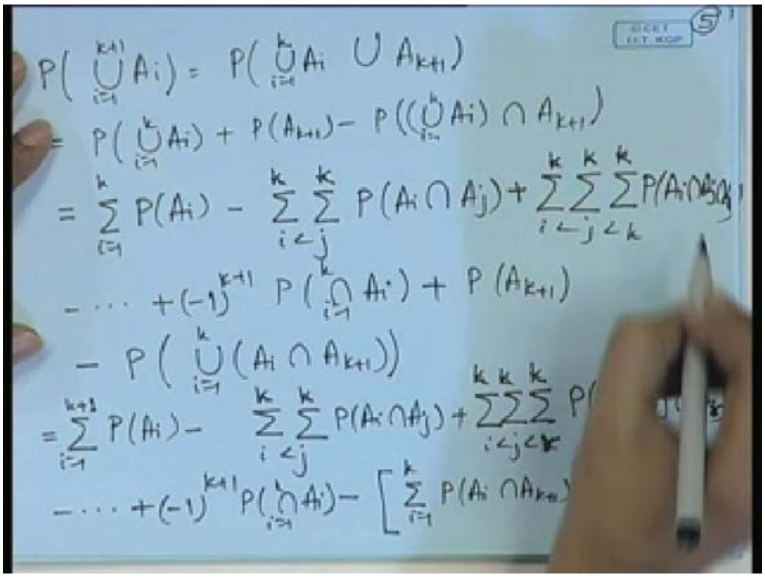
Now, let us look at the terms. Theterm sigma probability of AI intersection AK plus 1 can be combined with this term here where the minus get10g adjusted and therefore, if you see here now, we already had all the intersections up to K.Now, we have A 1 intersection AK plus 1 A2 intersection AK plus 1 and AK intersection A K plus 1. So, this gets adjusted here and you will get a term. So, the first term remains as such. Probabilityof AII is equal to 1 to k plus 1 in the second term you will get I less than J and now this summation is upto K plus 1 probability of AI intersection AJ.

Now, let us look at this term this term is AJ intersection AK plus 1 where I and Js are varying from 1 to k and if we look at the third term in the previous expression here we had all the intersections of 3 sets upto k. So, this term gets adjusted in this 1 and you will get plus triple summation probability of A I intersection A J intersection A l ARI less than J less than r upto k plus 1.

In A similar way, if I look at this term here it will be intersection of the 4 terms andthe last set is A k plus 1; that means, it is taking care of all the terms of the intersection taken 4 sets at A time. So, in this way all of the terms are combined if you look at this term this is actually intersection of all of the AIs from I is equal to 1 to k plus 1 and since there is a minus sign outside of the square bracket this becomes minus 1 to the power k plus 2. So, you will getminus 1 to the power k plus 2 probability of intersection A I I is equal to 1 to K plus 1 hence the result is true for all positive integral values of n.

Let us look at the some applications of this one. Now, before giving the application let me also consider thelimit of the probabilities or probability of the limitas I mentioned thatwe have defined monotonic sequences and for the monotonic sequences of the sets, the limit always exists. So, we have the following result for monotonic sequences of the sets.

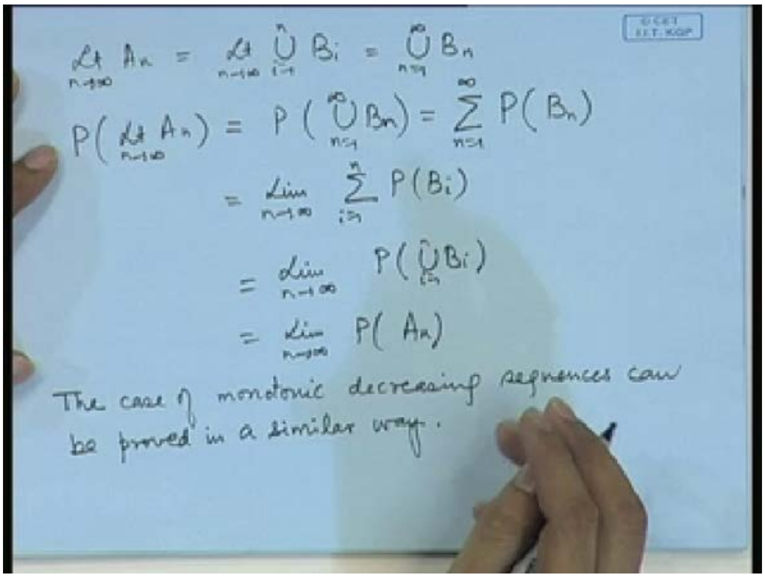
(Refer Slide Time: 22:28)



We have the following theorem if A nisA monotonic sequence of sets in B then, probabilityof limit of A n is equal to limit of probability of A n.

Toprove this result, let us consider A n to be say monotonically increasing sequence. Let AN be monotonically increasing sequence if that is. So, then limit of the sequence AN will become union of AN,Nis equal to 1 to infinity. Inorder to prove that we have to look at probability of limit - means probability of the union. Now, what we do? Wedecompose this union by defining a new sequence of sets by saying say B1 is equal to A 1; B2 is equal to A2 minus A1;B n is equal to A n minus A n minus 1, for n greater than or equal to 2. Ifwe look at this one basically, what we have done? Thesequence of sets is like this.Thisis say A1; this is A2; this is A3 and so on. So, if I look at the union of AIs, we are decomposing it as A disjoint union; this A1A2 minus A1 will be this portion then A3 minus A2 will be this portion. So, we will have that b n is a disjoint sequence of sets and A n is equal to union of BI from 1 to n.Naturally, this implies that probability of A n is equal to probability of BI sigma I is equal to 1to n.

(Refer Slide Time: 25:39)

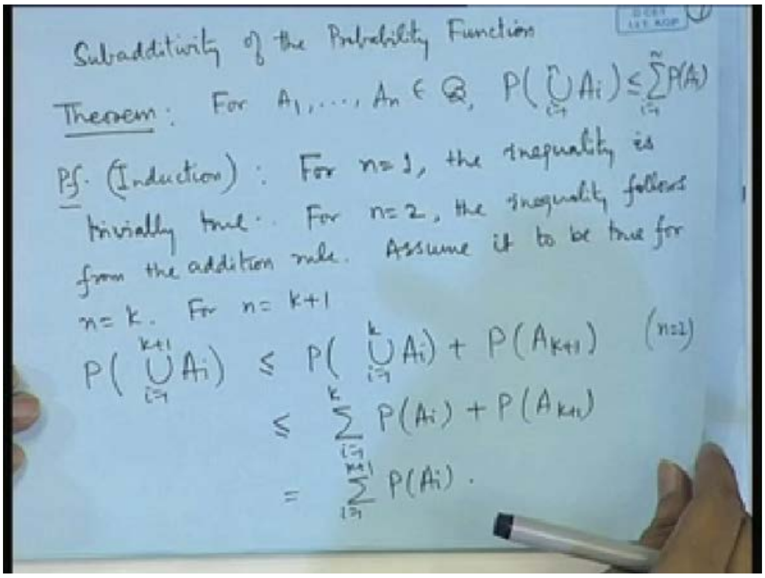


Now, if we look at limit of the sequence A n as n tends to infinity then, it is equal to limit of union BII is equal to 1to n, n tending to infinity which is equal to union of B n; n is equal to 1 to infinity because union BI is a monotonic increasing sequence andthe limit will be the ultimate union of these sets. So, if I look at probability of limit of A n as n tends to infinity then it is equal to probability of union B n n is equal to 1 to infinity. Now,B n is a disjoint sequence of sets; then by the axiom of the countable additivity this becomes probability of sigma probability of B n n is equal to 1 to infinity.

Now, this we can write as limit as n tends to infinity sigma I is equal to 1to n probability of BI which we can write as probability of union of BII is equal to 1to n which is equal to limit as n tends to infinity probability of A n. So, thus we have proved that probability ofA limit of A sequence ofmonotonic sequence of sets is equal to limit of the probability of the sequence of the sets. Wemay also consider the case of monotonically decreasing; now, that can be obtained by taking the complementations here or you can define in a reverse way.

The probability function we have assuming to be countably additive, but countably additiveaxiom implies that if we have a disjoint sequence then the probability of union is equal to the sum of the probabilities what if we do not have disjoint sequence for example, if I have 2 sets say A and B then we have probability of A union B is equal to probability A plus probability B minus probability of Aintersection B.So, that means if I remove probability of A intersection B from there then we get probability of A union B less than or equal to probability A plus probability of B this is called subadditivity.

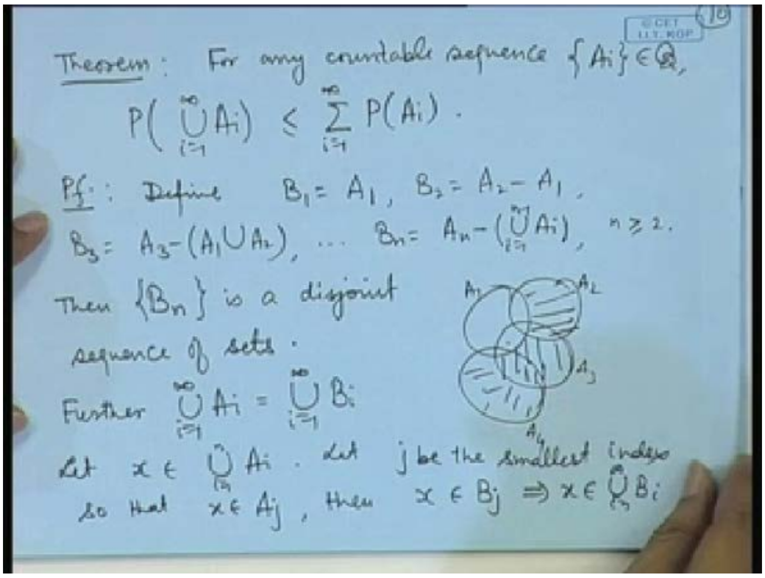
(Refer Slide Time: 28:50)



So, if in general if we consider any sequence of sets then the probability of union will be less than or equal to the sum of the probabilities. So, we have subadditivity of the probability functionand we can state it in the form of a theorem for A1A2A n belonging to B probability of union AII is equal to 1to n is less than or equal to sigma probability of A II is equal to 1to n. So, 1 can prove this by induction because for n is equal to 1 the result is true and if we look at for n is equal to 2 it is already shown to be true. So, for n is equal to 1 the inequality is triviallytrue for n is equal to 2 which we will require for extension from K to K plus 1 case. So, for n is equal to 2 the inequality follows from the addition rule.

So, assume it to be true for say n is equal K now for n is equal to k plus 1 we can write probability of union A II is equal to 1 to k plus 1 as less than or equal to probability of union A II is equal to 1 to k plus probability of AK plus 1, by using the result for n is equal to 2. So, now, on this we can make the use of assumption that upto n is equal to K - it is true. So, it becomes less than or equal to probability of A IIis equal to 1 to K plus probability of AK plus 1 which is nothing, but the sum of the probabilities I is equal to 1 to K plus 1therefore, by induction the result is true for all of them.

(Refer Slide Time: 31:45)

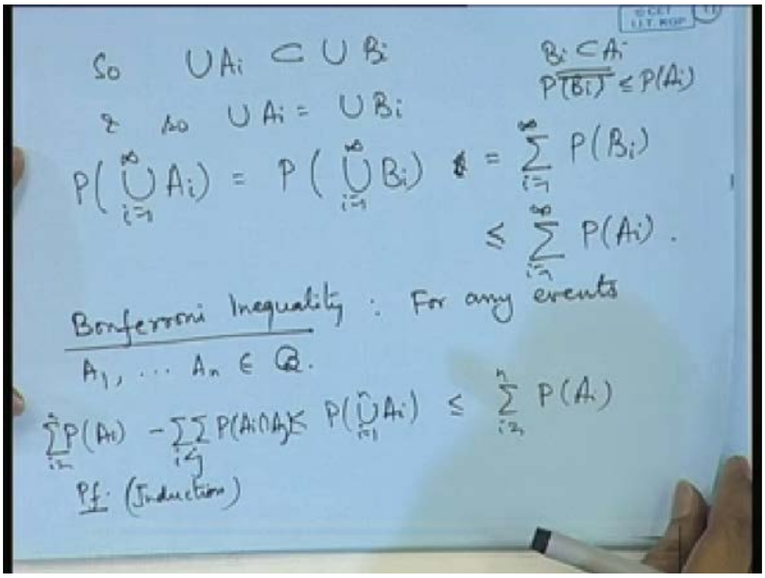


Now, if we want to prove the result forA countable number of these then, we can consider the decomposition. So, if we have for any countable sequence sayA I in bprobability of union A I Iis equal to 1 to infinity is less than or equal to sigma probability of AI I is equal to 1 to infinity in order to prove this one.I may consider the decomposition of union A I into A disjoint decomposition in the following way. Letus define say B1 is equal to A1;B2 is equal to A 2 minus A1;B3 is equal to A3 minus a 1 union A2 and so on. In general B n is equal to A n minus union of AII is equal to 1to n minus 1.

If we consider A Venn diagram then, it will be clear that what sets we are defining. Suppose, these sets are A 1,A2,A 3,A4, etcetera, then A 1 and then A2 minus A1 is this set; then A3 minus A1 union A 2 becomes this set A4 minus A1 union A 2 union A 3 becomes this set. So, naturally you can see here that we are considering the union asA disjoint union. So, then B n is a disjoint sequence of setsfurther union of A II is equal to 1 to infinity is equal to union of B II is equal to 1 to infinity.

To prove this let us observe that union of BI is already A subset of union BI because each of the BI is a subset of the corresponding AIs. Now, any point of AI let us consider say X belonging to union of AI let J be the smallest index. So, that X belongs to AJ then X will belong to BJ consequently X will belong to union of BIsas A result sinceunion of BI is already A subset of unionof A I we are now get10g union of AI is A subset of union of A I.

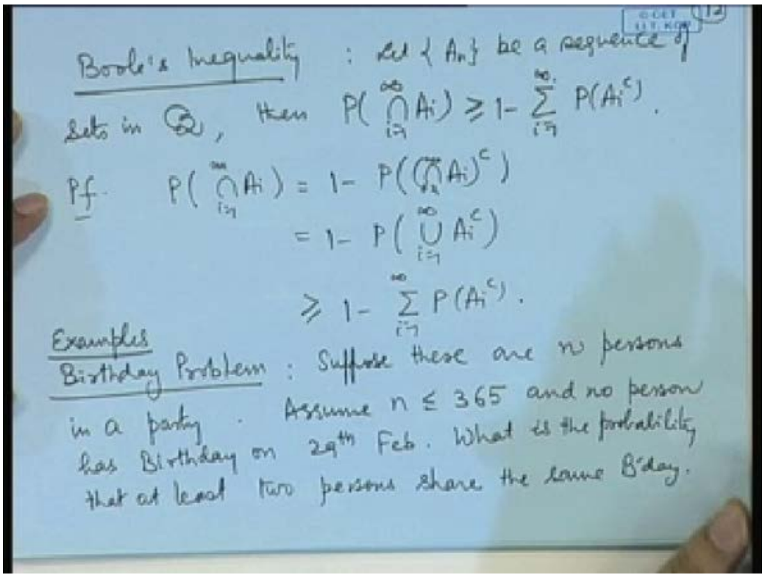
(Refer Slide Time: 35:16)



Therefore, we must have union of A I is equal to union of BIs. So, now if we consider probability of union of AII is equal to 1 to infinity it is probability of union of BI I is equal to 1 to infinity which is less than or equal to, which is actually equal to sum of the probability of BIs because,BIs are now disjoint and we can use the axiom of countable additivity. Now, each BI is A subset of AI. Therefore, probability of each BI is less than or equal toprobability of A I.Therefore, this becomes less than or equal to sigma probability of AIIis equal to 1 to infinity; this proves the countable additivity of the probability function.

We also have something called Bonferroni inequalities which basically give that the probability of the unions are bounded between 2 bounds. So, for any events A1A 2A n in B probability of the union which is already less than or equal to sum of the probabilities it is however, greater than or equal to probability ofminus. Letme not prove it here. Theproof will be by inductionwe can see the right hand side has already been proved. Toprove the left hand side, if we take n is equal to 1 then it is trivially true for n is equal to 2 there is equality by the addition rule. So, assuming for n is equal to K if we write for n is equal to K plus 1 then we can split it into 2 terms that is union of AII is equal to 1 to K union AK plus 1. Onthat we apply the addition rule and then apply the assumption for K that will prove the general Bonferroni inequality.

(Refer Slide Time: 38:01)

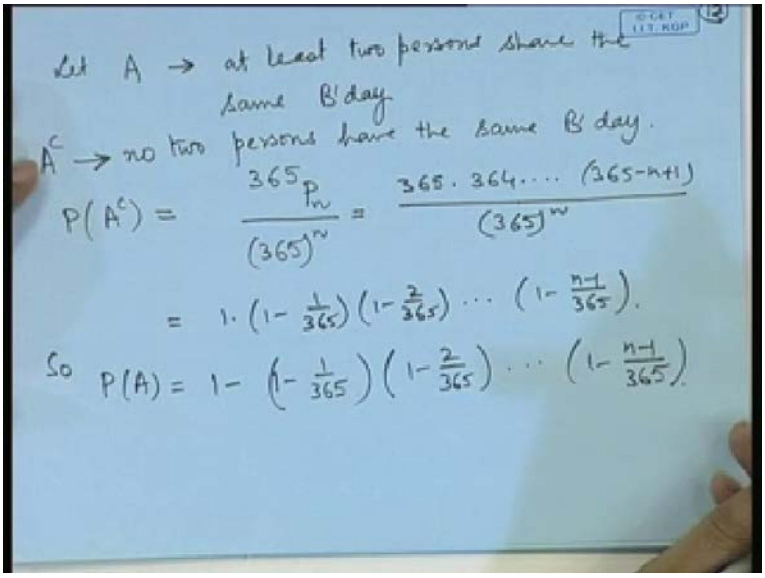


In a similar way we have what is known asBoole’s inequality the Boole’s inequality gives A relation between the intersection. Likewise for example, if I have A n is any sequence of sets in B then probability of intersection AII is equal to 1 to infinity is greater than or equal to 1 minus sigma probability of A I complement I is equal to 1 to infinity.

To prove this we simply use the subadditivity because we can write probability of intersection AIIis equal to 1 to infinity as 1 minus probability of intersection AI complement. Now, this can be written as 1 minus probability of union AI complement by using demorgan’S laws at this stage you can use the countable subadditivity. So, this will become greater than or equal to 1 minus sigma probability of A I complement.

Let me give some examples of applications of basic rules of probability let me start from a birthday problem supposethere are n persons in A party assuming that the number of persons is less than or equal to 365 and no person has birthday on 29 February. What is the probability that at least 2 personsshare the same birthday? Now, in order to analyze this problem, let us consider the set theoretic description.

(Refer Slide Time: 41:22)



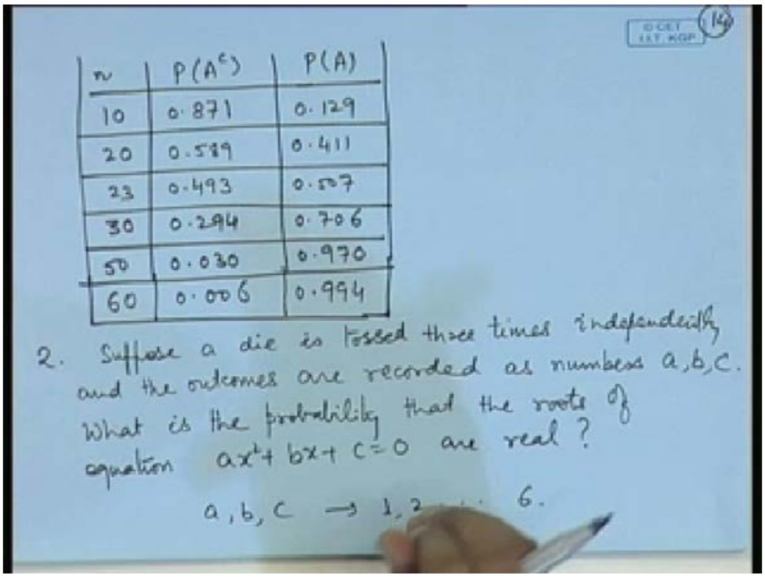
Let us consider A to be the event that at least 2 persons share the same birthday. Then, if you look at this event it is slightly complicated event in the sense that, 2 persons may share 3 persons may share and so on and finding out the probabilities of each of them may be a little bit complicatedbecause, if we say 2 persons share then which of the day it is and all others must be on some other dates and they should not be the same. Suppose, we say 3 persons share then which one of the 365 days and all other persons must be on distinct days which distinct days. So, this is a complicated way to analyze.

However, if we use theset theoretic representations, we can look at the complementary event. Acomplement - this means no 2 persons have the same birthday. Now, this becomes somewhat simpler because, if we look at the probability of A complement assuming all the birth dates to be equally likely this number will be simply 365P ndivided by 365 to the power n. Here, the denominator denotes the total number of possibilities for n persons to have birthdays because, each person can have any of the 365 days as a possible birthday and therefore, n persons can have possible number of birthdays as 365 to the power n.

If we make the assumption that n1 of them have the same birthday then it becomes a problem of choosing n numbers out of 365 which are distinct. So, it is nothing but the number of permutations taking n at a time from 365 that is equal to 365 into 364 upto 365 minus n plus 1 divided by 365 to the power n, which we may write as a way of representation as 1;1 minus 1 by 365; 1 minus 2 by 365 and so on upto 1 minus n minus

1 by 365 and so on.. So, probability of A becomes 1 minus the product given by these terms. An interesting thing; what we do? Lookat that how many people are required; so, that at least 2 will share a birthday. If we think from a layman point of view then, we may think that the numbers since the number of possible birthdays is 365 to the power n. So, n should be somewhat large in order that this probability is significant. So, let us look at the table of probabilities.

(Refer Slide Time: 45:03)



let us consider sayn probability of A complement and probability of A. So, A simple calculation table can be prepared if I have n is equal to 10; the probability of A complement is point 871 and consequently, probability of A becomes point 129. Ifwe take n is equal to 20, probability of A complement is point 589 and probability of A becomes point 411. Ifwe take n to be 23 then, probability of a complement is point 493 and probability of A becomes point 5 07. That means, with as less as only 23 persons the probability that at least 2 share a common birthday is more than 50 percent. So, it is from a layman's thinking; this is counter intuitive.

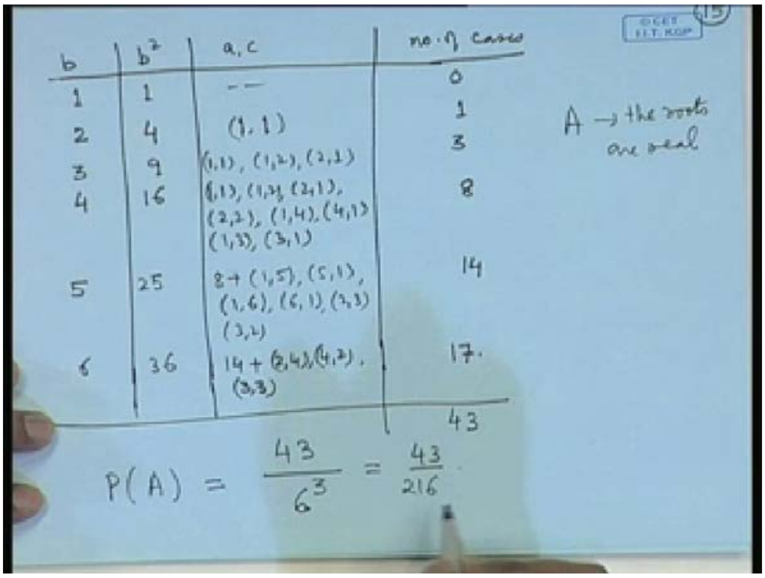
We need very few persons to at least 2 of them to share a common birthday. If I take n is equal to 30 then, this probability becomes point 706. Ifwe take 50, the probability is point 97 and for n is equal to 60 the probability is point 994; it is nearly 1. That means, in a set of 60 people, the probability is nearly 1 that at least 2 of them will share a common birthday.

So, here you can see that theelementary rules of probability have been used for calculation.For example, we have use the property of the complementation to evaluate the actual probability we have use the method of classical probability by assuming all the birth dates to be equally likely for all the persons.

Let us look at some other applications of the basic rules of the probability. Suppose, a die is tossed 3 times independentlyand the outcomes are recorded as numbers A B and C.What is the probability that the roots of equation A square X plus B, x plus C is equal to 0 are real.

Now, if we want to calculate this probability here the outcomes A B and C are random each of the values of ABand C can be numbers 12 upto 6.Therefore, the quadratic equation AX square plus BX plus C is equal to 0 will have the real roots if B square minus 4A C is positive. So, we have to look at the number of cases where b square minus 4AC is greater than or equal to 0.

(Refer Slide Time: 49:35)



So, this has to be done through an enumeration and we can prepare the table that what are the possibilities of B and therefore, the corresponding values of B square. Whatare the possible values of A and C which lead to 4AC being less than or equal to B square? So, let us take say B is equal to 1 then B squareis equal to 1; that means, there is no case which will give me 4 AC to be less than or equal to B square. So, there is no possibility here.

So, if we look at the number of cases this is 0 if we take b is equal to 2 then b square is equal to 4 and if I consider a andC to be 11 then 4AC will become 4. So, there is 1 casewhich will give me B square greater than or equal to 4AC.Ifwe consider B is equal to 3 then B square is equal to 9 now 111 2 and 2 1. Thereare 3 cases which will give me B square greater than or equal to 4AC. Ifwe have B is equal to 4 then B square is equal to 16, we will have the cases 111 2 2 1 2 2 which will correspond to 4. So, 14 4 1 13 3 1 1 23 4 5 67 8 cases are there which will give me B square greater than or equal to 4AC.

If we have B is equal to 5 then B square is equal to 25 then all the above cases that is 8 cases plus we will also have 15 5 1 andpossibly 1 4 4 1. So, we will also have 16 61 2 3 3 2, basically 1 3 3 4 5 6 more cases. So, 14 cases are there. If I have B is equal to 6 then,B square becomes 36 andall the 14 cases plus we will also have 2442 then 25 is not possible.33I think 33 must have come here itself because, no it will not come here 33 will come here because this will give me 9. So, there are 17 cases. So, if we look at the total number of cases it is 39, 42, 43 cases are there. Totalnumber of cases is 43 and the total number of possibilities if I define a to be the event that the roots are real then, the probability of that will be given by the favorable number of cases divided by the total number of cases which is 6 cube here because,3 dice - each of them have 6 possibilities. So, the total number of possibilities are 6 cube that is 43 by 216.

Likewise, in this problem we may also find out the probability of thequadratic equation to have complex roots or thereal roots, but equal, etcetera. Wemay consider all types of possibilities. So, I will end today’s lecture by this; thank you.