

so i had already initiated the discussion on the conditional probability in my last lecture ah you can see what i will recapitulate if i am considering say tossing of a fair die and i consider the event a as 1 occurs then what is the probability of it it is 1 by 6 now i modify my statement what is the probability that one occurs given that an odd number occurs and i call it probability of a given that b has occurred now out of these six possible outcomes in the fair die there are three outcomes which are odd numbers one three and five

So now out of this if i say what is the probability that one has occurred then it will be one by three

So this is actually leading to the concept of conditional probability that means if i have additional knowledge of certain event in an experiment then my original probability may get modified

So let us define let a and b be any two events where the probability of b is positive ok then the conditional probability of a given that event b has occurred

So this is defined in the following fashion we write this notation probability of a conditioned on b

So it is read as probability of a given b ok it is defined to be probability of a intersection b divided by probability of b if i apply this formula to this particular case

So you can see probability of a intersection b

So a is one occurs and b is one three five

So a intersection b is one

So it will become one by six and probability of b is half

So one by six divided by one by two that will give me one by three

So this definition is conforming to this original understanding of what is conditional probability ah

So let me extend this argument little bit let me call this relation number one from relation one we can write probability of a intersection b is equal to probability of b into probability of a given b similarly if probability of a is positive then we can define probability of b given a also

So that will become probability of b intersection a which is same as a intersection b divided by probability of a

So this will lead to the statement probability of a intersection b is equal to probability of a into probability of b given a let us look at these statements two and statements three and let us talk in terms of probabilistic interpretation statement number two on the left hand side i have probability of simultaneous occurrence of two events a and b

So i am saying it is equal to probability of one of the events multiplied by probability of second event given the first event

So the simultaneous occurrence probability can be calculated as a product of two probabilities one is conditional and another is probability of one of the events which is also called marginal probability and in this statement the roles of a and b have been interchanged this rules ah one and this two and three these are called multiplication rules

So the basic concept of multiplication rule is that probability of simultaneous occurrence of two events can be calculated as a product of two probabilities

So multiplication is there

So that is why it is called a multiplication rule now immediately the idea comes that can i extend it to three events the answer is yes in fact i can write this multiplication rule for n number of events and the proof is again by induction as you have seen in the case of addition rule

So let me give that now the general multiplication rule

So let us consider say a one a two a n b events and in order to define the conditional probability ah the event which is occurring in the denominator that is like in the denominator i put probability of a or probability of b then their probabilities must be positive otherwise the ratio will not be defined therefore i can put a condition probability of say intersection of a i i is equal to one to n is positive ah now this actually condition is enough to say that probability of a one will be positive probability of a two will be positive or probability of a n will be positive because this set is actually the smallest set when i consider several ah events and i take intersection of all of them then that is the smallest set

So if I put the intersection of the smallest set to be having a positive probability then all of the corresponding events that means which will be occurring as individuals or if they are occurring as taking intersecting two at a time like a one intersection a two a three intersection a four etcetera or taking intersection three at a time all of them will have positive probabilities

So let me just write here comment this will ensure that all conditional probabilities are well defined

So the general addition rule is multiplication rule is the probability of intersection a  $i$   $i$  is equal to 1 to  $n$  that is equal to probability of a 1 into probability of a 2 given a 1 into probability of a three given a one intersection a two and

So on probability of a  $n$  given a  $i$   $i$  is equal to one to  $n$  minus one

So this is the general multiplication rule the proof is again using the principle of mathematical induction suppose I call this equation number four the relation four can be proved using the principle of mathematical induction

So for example I can say for  $n$  is equal to one the statement is always true as it reduces to the left hand side for  $n$  is equal to one it will give me only one term that is probability of a one similarly the right hand side also when  $n$  is equal to one will have only one term

So this statement is probability of a one is equal to probability of a one now assume it to be true for  $n$  is equal to  $k$  then for  $n$  is equal to  $k$  plus one the statement is like this probability of intersection a  $i$   $i$  is equal to one to  $k$  plus one

So firstly I consider it for splitting into two parts a one intersection a two intersection a  $j$   $j$  is equal to three to  $k$  plus one

So this is becoming probability of a one intersection a two into probability of a  $j$  sorry intersection a  $j$   $j$  is equal to three two  $k$  plus one given a one intersection a two now here I can again apply for  $n$  is equal to  $k$

So this can be further expanded and this I can write as probability of a one into probability of a two given a one and

So on let me consider one example here let  $a$  be the event that a person gets cured from a disease and  $b$  is the event that person gets some treatment some medical treatment suppose I am assuming that the probability of  $b$  is 0.9 that is 90 percent of the people have access to the medical treatment which are suffering from this disease and 80 percent of the persons who actually get the treatment get cured then what is the probability of a intersection  $b$  then it is equal to probability of  $b$  into probability of  $a$  given  $b$  that is point nine into point eight that means point seven two

So you can give the statement that 72 percent of the persons who get the disease actually get cured because 90 percent of the people have access to the medical treatment and out of the people who actually get the treatment 80 percent get cured

So overall 72 percent people actually get cured

So this is a direct application of the multiplication rule here now this multiplication rule and the conditional probability concept is useful to consider a calculation of probabilities where a certain event may arise out of several things

So this is called something like a cause effect relationship

So for a particular effect there can be several causes

So for example a person dying

So the cause of the death the cause of death can be due to disease due to accident or due to natural causes etcetera

So when we are calculating probability of the final effect we may have to take into account various causes

So each of the causes now in probability this concept has been formalized by the concept of total probability

So let me give this here theorem of total probability let  $v$  one  $b$  two let me write a finite number of events be pairwise disjoint and exhaustive events such that probability of  $b$   $i$  is positive for all  $i$  ok this is the notation for that for each of the  $i$  that is probability of  $b$  1 is positive probability of  $b$  2 is positive probability of  $b$   $n$  is positive then for any event  $a$  probability of  $a$  can be written as probability of  $a$  given  $b$  1 into probability of  $b$  one plus probability of  $a$  given  $b$  two into probability of  $b$  two and

So on a given  $b$   $n$  into probability of  $b$   $n$  let us look at the proof of this

So ah we have here exhaustive events that means  $s$  is equal to union of  $b$   $i$   $i$  is equal to one to  $n$  where  $s$  is the sample space

So we can write then  $a$  as a intersection  $s$  why because any event will be a subset of sample space

So if i take a intersection  $s$  i will get only a now the advantage of this is that i can write  $s$  as union of  $b_i$   $i$  is equal to 1 to  $n$  now this is intersection union

So i can apply the distributive law

So the distributive law will give me union of  $i$  is equal to one to  $n$  a intersection  $b_i$  now you see what i have done there is an event  $a$  which i have expressed as union of certain sets and what type of sets

So let us look at this also i have assumed here that  $b_1$   $b_2$  etcetera they are disjoint sets

So let us consider an rough diagram here suppose this is my sample space ok and i have event say this is say my event  $b_1$  this is event say  $b_2$  this is event say  $b_3$  and

So on and say suppose this is event  $b_n$

So i had intentionally designed in such a way that these are disjoint as well as the union of all of them is actually equal to  $s$  now  $a$  is any event here ok

So  $a$  is some event

So what will happen to a intersection  $b_1$  a intersection  $b_1$  is this a intersection  $b_2$  is this dotted part a intersection  $b_3$  is say ah let me draw this curly lines here a intersection  $b_n$  is say suppose i put here ah rounded figures here ok

So this is a intersection ah sorry this part will be a intersection  $b_n$  not this now you see if  $b_1$   $b_2$   $b_3$   $b_n$  etcetera they are disjoint then a intersection  $b_1$  a intersection  $b_2$  a intersection  $b_3$  a intersection  $b_n$  are also disjoint since  $b_1$   $b_2$   $b_n$  they are pair wise disjoint the sets are events a intersection  $b_1$  a intersection  $b_2$  and

So on a intersection  $b_n$  will also be pairwise disjoint because they are actually a intersection  $b_1$  is actually a subset of see this is a subset of  $b_1$  a intersection  $b_2$  is a subset of  $b_2$  and

So on a intersection  $b_n$  is a subset of  $b_n$

So if this  $b_1$   $b_2$   $b_n$  etcetera are disjoint then a intersection  $b_1$  a intersection  $b_2$  a intersection  $b_n$  will also be disjoint now what we have done  $a$  has been written as a union of disjoint events if these are disjoint then by the first ah i have given you the axiom of additivity that is the kalmogorov's axiom there the third axiom was that if you have pairwise disjoint events then the probability of the union is equal to some of the probabilities

So if i apply that by the axiom of additivity we get probability of  $a$  is equal to probability of union a intersection  $b_i$   $i$  is equal to one to  $n$  that is equal to summation  $i$  is equal to one to  $n$  probability of a intersection  $b_i$  now this a intersection  $b_i$  again i can apply the multiplication rule on this

So i will get it as summation  $i$  is equal to one to  $n$  probability of a given  $b_i$  into probability of  $b_i$  which was actually the original statement of the theorem of total probability that is the probability of a is probability of a given  $b_1$  into probability of  $b_1$  plus probability of a given  $b_2$  into probability of  $b_2$  and

So on probability of a given  $b_n$  into probability of  $b_n$

So that statement we have proved here ah to look at it into the context actually it means that if a event can be decomposed as a union of  $n$  mutually exclusive events then the probability of the final event is in terms of conditional probabilities of given each of them and the marginal probability of each of those events

So let me give one example here a fair coin is tossed once if a head comes up a fair die is tossed once and if a tail comes up a fair die is tossed twice we want to find the probability that at least a six is observed ok

So let us apply the theorem of total probability here because there are two possibilities you can have a head or you can have tail

So if i consider the event let  $h$  denote the event that head comes up and say  $t$  denotes the event that tail comes up that means when a coin is tossed then you may get a head or a tail and i use this notation  $h$  and  $t$  for denoting that and  $a$  is the event that a six is observed

So then probability of  $a$  can be written as probability of a given  $h$  into probability of  $h$  plus probability of a given  $t$  into probability of  $t$  because  $h$  union  $t$  is the full sample space here

So you are able to get you can say the theorem of total probability here by is splitting into events  $h$  and  $t$  now let us look at the probability here if a head is observed then

the consequences that a die is tossed once

So die is tossed once what is the probability of head here what is the probability of a six that will become one by six and what is the probability of a head here that is half here because it is a fair coin in the second part a tail comes then a fair die is tossed twice what is the probability of at least a six

So this i can calculate using that there is no six

So no six becomes in one toss probability that we do not get a six the probability will be five by six now two times it is done and both the times i do not get a six then the probability will be five by six square

So if i take one minus that this is meaning that we get at least one six and the probability of a tail is half now this we can easily simplify and we get seventeen by seventy two

So you are getting and let us also try to numerically interpret it seventeen by seventy two is ah you can say little bit higher than point five ah sorry ah it is little bit higher than ah fourteen by seventy that is point two when you have toss of a die you have the probability of a six as one by six ok ah

So that is roughly 16 percent here but here you are getting much higher than that seventeen by seventy two

So the reason is that we are also allowing the possibility of the fair die to be tossed twice therefore the probability of getting a six has increased from one by six to seventeen by seventy two actually one by six is twelve by seventy two and here you are getting seventeen by seventy two

So numerically you can see by adding additional die only in part we have added the probability here this is an application of the theorem of total probability now in the theorem of total probability what we are doing is that when an event is there in order to calculate the probability of that event we are looking conditioning on several causes

So i as i mentioned to you cause effect relationship kind of thing now we can further extend this idea suppose the effect is available to you that means you already observe something and but we did not notice what was the cause and then we go back and see whether we can calculate the probability of the original cause what could have been the cause

So if we do that then we are looking at the probability in a reverse way or you can say conditional probability with reverse conditioning here we have given the concept that probability of a given b one probability of a given b two etcetera but if i know what is a then what is the probability of b one

So that means i am calculating probability of b one given a probability of b two given a etcetera

So this concept is formalized in the famous statement of bayes theorem ah it is named after thomas base and this book was published on 1763 after his death but the statement continues and this bayesian theory has become very popular now in the context of decision theory in various rules of estimation base estimators are used bayesian tests are used

So bayesian decision rules are used in practice

So let me give the statement here

So let  $b_1, b_2, \dots, b_n$  be any events ah

So again we are taking disjoint events which are exhaustive and probability of each of them is positive let a be any event with probability of a positive then probability of say  $b_r$  given a that is probability of a given  $b_r$  into probability of  $b_r$  divided by  $\sum_{i=1}^n$  probability of a given  $b_i$  into probability of  $b_i$  i is equal to one to n this statement is ah the famous statement of the bayes theorem ah let us look at the proof and as again you can see since we have actually restricted our attention to the set theoretic notations and the exomatic approach the proofs are actually very simple

So probability of b are given a this is statement is true for r is equal to 1 2 and n that means for any event i can calculate it

So we apply the definition of the conditional probability

So it becomes probability of  $b_r$  intersection a divided by probability of a and on this if you look at the numerator i can apply the multiplication rule in the reverse way that means i can write it as probability of a given  $b_r$  into probability of  $b_r$  this is using multiplication rule and in the denominator this probability of a i simply apply the theorem of total probability because i am having the same conditions that the events are disjoint and exhaustive then for designed and exhaustive events probability of a is equal to  $\sum_{i=1}^n$  probability of a given  $b_i$  probability of  $b_i$

So I can write that here in the denominator using theorem of total probability  
So actually the statement has been proved just in two steps here one step I just wrote the definition of conditional probability in the second step in the numerator I applied the multiplication rule and in the denominator I have applied the theorem of total probability

So the base theorem has been established here

So let us just consider one simple application here

So a say computer manufacturer procures chips from three suppliers say  $b_1$ ,  $b_2$ ,  $b_3$  in proportion say two by five three by ten and say three by ten respectively  
So out of his total procurement of chips  $A$  for fixing in the computers he purchases two by five proportion from  $b_1$  three by ten from  $b_2$  and three by ten from  $b_3$   $A$  it is known from experience that one percent of chips from  $b_1$  are defective five percent from  $b_2$  are defective and ten percent from  $b_3$  are defective ok  
So a chip is randomly selected from the collection of the manufacturer it is found to be defective

So what is the probability that it was supplied by  $b_1$  now you see here this is cause effect relationship because the final effect whether the chip is defective or not can be cast by the supply of  $b_1$ ,  $b_2$  or  $b_3$  now we know finally the effect that it is actually defective

So we are looking at what was the cause that means who caused it

So that is whether it was  $b_1$  or  $b_2$  or  $b_3$  what are the respective  $A$  probabilities here

So let us calculate it here

So if we apply the Bayes theorem then we will have what is the probability of  $A$

So let me let me define the events here suppose I define the event  $A$  as chip is defective  
So what is the probability of  $A$  the probability of  $A$  will be summation probability of  $A$  given  $b_i$  into probability of  $b_i$   $i$  is equal to one two three where  $b_i$  denotes the event that the chip is supplied by  $b_i$  manufacturer

So here I know all these things because probability of  $b_1$  is known to me two by five probability of  $b_2$  is three by ten probability of  $b_3$  is three by ten similarly probability of  $A$  given  $b_1$  is one by hundred probability of  $A$  given  $b_2$  is five by hundred probability of  $A$  given  $b_3$  is ten by hundred because we know the probability of defective chips from each of the suppliers

So if we substitute these values here I get here  $\frac{1}{100} \times \frac{2}{5} + \frac{5}{100} + \frac{3}{100} \times \frac{3}{10}$

So that is equal to we can calculate this easily that is equal to point zero four nine  $A$  you can also try to appreciate this number that means basically we are saying approximately five percent of the chips are defective which the person procures the manufacturer procures see he is procuring from three one of them has one percent defective another has five percent defective and another is ten percent effective now he is taking in miscellaneous quantities from each of them overall his collection will have nearly 0.49 that is you can say nearly 5 percent effective chips now if I want to calculate the probability of say  $b_1$  given  $A$  then by the base theorem it is becoming probability of  $A$  given  $b_1$  into probability of  $b_1$  divided by probability of  $A$

So that is equal to  $\frac{1}{100} \times \frac{2}{5} \div 0.049$

So after simplification it turns out to be simply  $\frac{4}{49}$  similarly I can calculate probability of  $b_2$  given  $A$  that is equal to probability of  $A$  given  $b_2$  into probability of  $b_2$  divided by probability of  $A$

So that is equal to  $\frac{5}{100} \div 0.049$  that is equal to  $\frac{15}{49}$  and probability of  $b_3$  given  $A$  is equal to probability of  $A$  given  $b_3$  into probability of  $b_3$  divided by probability of  $A$  that is equal to  $\frac{3}{100} \div 0.049$  that is equal to  $\frac{30}{49}$  let us see the effect of the numbers here see what were the probabilities of  $b_1$  that was two by five probability of  $b_2$  was three by ten the probability of  $b_3$  was also three by ten but the conditional probabilities now have become quite different from two by five which is approximately forty percent it has gone down to four by forty nine that is less than ten percent actually it has become less than ten percent whereas this number has appreciated from three by ten it has become fifteen by forty nine and this number from three by ten it has become thirty by forty nine which is nearly sixty percent what does it mean since we are actually we have observed the final effect that means the chip is actually defective

So which company is more likely to supply the defective chip because we are purchasing

things from b 1 b 2 and b 3 and b 3 is having maximum number of defective that is 10 percent defective chips are there

So if the chip is defective it is highly possible that it was supplied by b three because the almost sixty percent chance is that from b three that defective chip was supplied and somewhat less chance from b two that is five percent effective

So that has become now more than five percent here and the share of b one has become much less although we are taking forty percent product from b one but since the number of defectives from b one is much less that is one percent only therefore if ultimately the chip is defective the chances are less that it would have been supplied from b one

So basically this is the effect of the ah base theorem or you can say because we are able to do the cause effect relationship in perspective this has this can have applications in several areas especially in crime detection doing some forensic examination etcetera here we actually know the final thing and then we want to detect the causes

So this particular result particularly became very famous and currently it is being applied in various areas of science and engineering now after this we have considered the concept of conditioning that means conditioning actually affects occurrence of the probability of occurrence of certain even what if it does not affect

So if it does not affect we call it to be independent events

So following from the conditioning actually we can give a new concept that is called independence of events

So now we can say that if event b has no effect on the probability of occurrence of a then it will mean that probability of a and probability of a given b must be same

So if occurrence of b does not affect the probability of a then this statement will be true provided of course this is well defined that means i am assuming that ah probability of b is positive

So this is statement you can write like this probability of a intersection b divided by probability of b that is equal to probability of a which you can further write as probability of a intersection b is equal to probability of a into probability of b now if you see this last statement this is a symmetrical statement in this statement actually i am putting b as the conditioning event but here there is no distinction between a and b because i am simply saying probability of the simultaneous occurrence of a and b is equal to the product of the individual probabilities of a and b in fact if i consider another one suppose i consider if occurrence of a does not affect the probability of b then what will be the statement the statement will be like this probability of b given a is equal to probability of b again if i simplify this i can write it as probability of b intersection a divided by probability of a is equal to probability of b which is again equal to i take it to this side

So it becomes probability of b intersection a is equal to probability of b into probability of a

So if you look at event ah this is statement number one and statement number two these 2 statements are actually same and the difference between this statement this statement and this says that these statements are not symmetric here conditioning is on a here conditioning is on b but if you look at this final consequence this statement is symmetric

So we consider this as the definition of independence of two events that means we say events a and b are independent if this condition one is satisfied because this will mean that occurrence of a does not affect occurrence of b occurrence of b does not affect the occurrence of a

So basically that should be the physical understanding of the concept of independence

So we define like that

So we define events a and b to be independent if probability of a intersection b is equal to probability of a into probability of b

So let me give a very simple example suppose two dice two fare dyes are tossed simultaneously ok and i consider the event let a be the event that even number on the first die and suppose i consider b to be the event even number on the second die ok i want to check whether a and b are independent

So what is the probability of a see if we look at the sample space when we toss two fare dice then it is written like this i will have numbers 1 1 1 2 and

So on 1 6 2 1 2 2 2 6 and

So on finally you will have 6 1 6 2 and

So on 6 6.

So in how many of this there is an even number on the first die

So if you look at the first die here in the second row when two is occurring then even number on the first

So there are six such cases similarly there will be four one four two four six six one six two six six that means there are total eighteen cases out of thirty six cases where there is an even number on the first die

So the probability of that becomes eighteen by thirty six that is equal to half similarly if i consider what is the probability of b that is an even number on the second die

So if you look at the second die second i the numbers if you look at the second column here one two two two up to six two

So here on the second eye you have an even number similarly if you have one four two four up to six four or one six two six six six etcetera there is again eighteen cases where you have an even number on the second die

So the probability of that becomes half now let us look at what is the probability of a intersection b now a intersection b means that there is an even number on the first die and there is an even number on the second i

So what are the cases

So in the second row if you look at you have two four ah two two two four and two six there are three cases in the second row in the fourth row you will have four two four four and four six and in the third sixth row you have six two six four and six six there are total nine cases where you have an even number on the first die and an even number on the second eye

So the probability becomes nine by thirty six that is equal to one by four

So here you can easily observe that probability of a intersection b is equal to probability of a into probability of b

So here you can say events a and b are independent events a and b they are independent now naturally you can think of extending the concept to more than two events now

naturally it comes out that if i consider three events then what will be the conditions probability of a intersection b is the proper product of probability a and probability b probability of b intersection c probability of a intersection c but at the same time you have to take all the three also

So if we define we say events a b c are mutually independent if probability of a intersection b is equal to probability of a into probability of b probability of b intersection c is equal to probability of b into probability of c probability of c intersection a is equal to probability of c into probability of a and probability of a intersection b intersection c is equal to probability of a into probability of b into probability of c actually these three conditions they are called conditions of pairwise independence and if you take all the four then this is called mutually independent in fact there can be cases where out of these four conditions may be three conditions are satisfied or two conditions are satisfied

So all the conditions may not be satisfied in that case we will not say that the events are independent in the previous example let c be the event that the sum is even ok if we say the sum is even then what are the possibilities here what is the probability of c first of all

So the sum is even in 1 1 1 3 1 5 2 2 2 4 2 6 3 1 3 three three five four two four four four six five one five three five five six two six four six six

So there are again total eighteen cases

So you will get probability of c is equal to half if i consider probability of a intersection b that is the first one is even and the sum is even if you consider that then out of total eighteen cases you will get only nine cases because you will get two two two four two six etcetera four two four four four six six two six four six six

So total nine cases will be there

So you will get one by four ah sorry a intersection c because the intersection b we have already calculated and if we consider probability of b intersection c that is the second one is even and the sum is even then it means the first one has to be even then this becomes 9 by 36 again the cases which i enumerated just now

So a b c are pairwise independent but if i look at what is the probability of a intersection b intersection c that means all the three are true then all the three are true again in nine cases only

So this probability is also one by four

So it is not equal to probability of a into probability of b into probability of c

So  $a$ ,  $b$ ,  $c$  are not mutually independent in this particular example. I have considered three events where  $a$ ,  $b$ ,  $c$  they are  $a$  and  $b$  are independent,  $b$  and  $c$  are independent,  $a$  and  $c$  are independent but if I take all the three together they are not independent. The reason is that if I take  $a$  and  $b$  then  $c$  is automatically certified to be even because if  $a$  is even,  $b$  is even then  $c$  is dependent upon that. That is why, ah, this is not independent of  $a$  and  $b$ . Ah, in the next class I will consider various problems on probability: various kind of applications of addition rule, multiplication rule, the theorem of total probability, Bayes theorem, and the concept of independence. You

Prutor@elitk