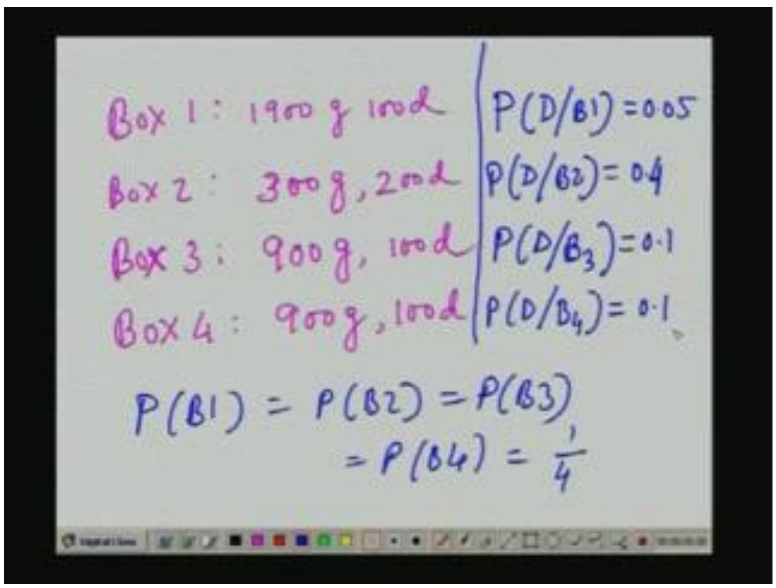
**Probability & Random Variables**
**Prof. M. Chakraborty**
**Department of Electronics & Electrical Communication Engineering**
**Indian Institute of Technology, Kharagpur**

**Lecture - 4**
**Introduction to Random Variables**

So, in the previous class I gave a small problem for you to work out. So, I am not sure whether you have done it or not, but I would like to spend some time now to solve, solve that problem for your sake.

(Refer Slide Time: 01:12)




Let me quickly repeat the problem that I gave which said that there is box there are 4 boxes. Box 1: it consists of 2000 item out of which 1900 items were good and 100 were defective. Box 2: it had 3000 good items 200 defective items. Box 3: it had 900 good items 100 defective items and last box, box 4 it also had 900 good items and 100 defective items.

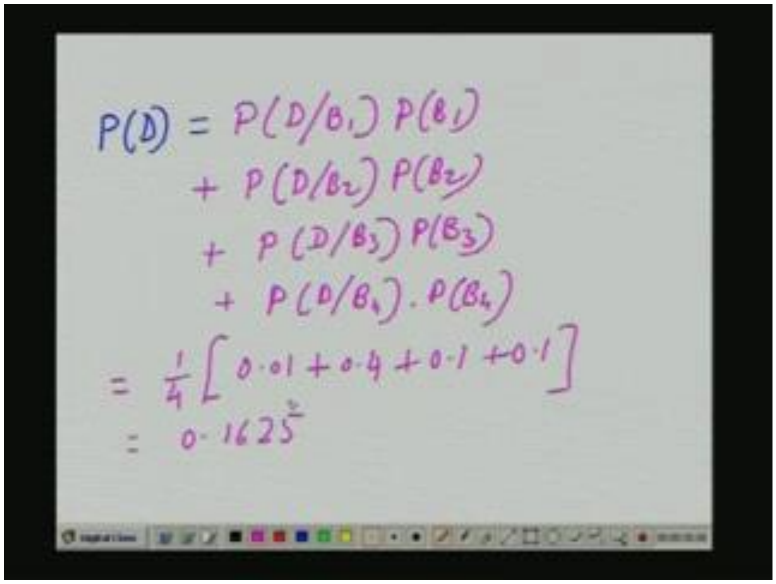
Question that was asked was that I mean we select any box in particular box at random first and then from that box we remove just only 1 element. Where first question is, what is the probability that, the particular element which is being removed is a defective 1? So, to answer that first we see that each of the boxes see the boxes they are chosen purely at random and there is no particular preference for a special box then box 1 box 2 box 3 or box 4.

The probability that I choose box one. So, B 1 or B 2 or B 3 or B 4 that is 1 by 4, because is uniform; that means, P B 1 is same as P B 2 is same as P B 3 is same as P B 4 is equal to 1 by 4 the uniform. This is very safe assumption, because there is to preference for a particular box. And then right across I can write that suppose I choose box 1, then what is the probability that the particular item that is removed from that box is a defective 1.

That is P D by B 1 this is equal to now there are 2 200 elements out of which only hundred are defective right. So, probability that the particular ball particular item which has been removed is having its defective is having a probability 100 by the total that is 2000 which means 0.05. Similarly, PD by B 2 that is assuming that, box 2 has been selected condition to the fact that B 2 has been selected. What is the probability that the element that is removed item that is removed is a defective 1?

So, in this case total number of items is 5 hundred and 2 hundred are the defective once. So, probabilities 200 by 500 that is 0.4. In the same manner PD by B 4 sorry this is B 3 just a minute please. Now, here total is I mean in total we have thousand elements that hundred are defective. So, this will be just 100 by 1000 which is 0.1 and this is identical to this. So, PD by B 4 this is again point 1. So, the question is what is the probability that the item that is removed is defective?

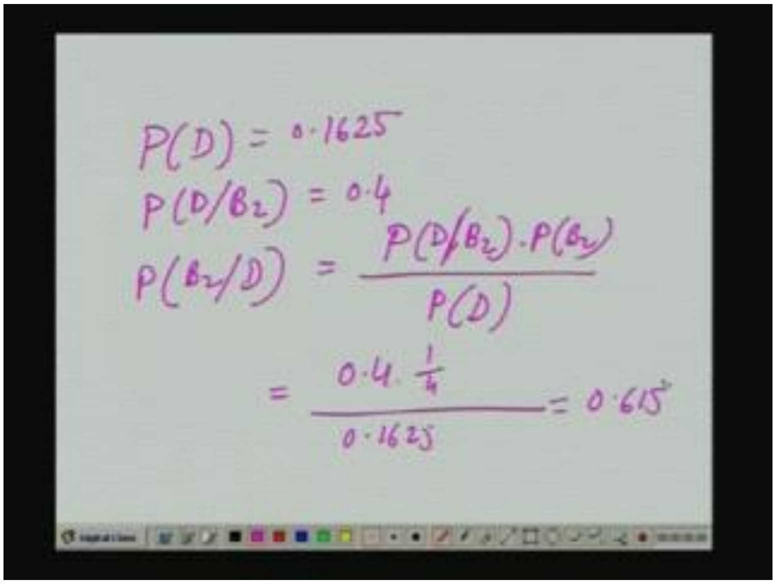
(Refer Slide Time: 06:08)




So, we will apply this total probability theorem that is probability that the a selected item is defective is given by … That is assuming B 1 is selected the probability is PB 1 then condition to that the probability, there is defective is P of B by B 1 plus PD by B 2. And we do not know all these values these are all 1 by 4, you can take 1 by 4 common. And these individual probabilities are 0.01 plus I think 0.4 plus 0.1 plus 0.1.

That breaks it. In fact, if you calculate it will turn out to be. So, this is the probability that the selected item is a defective 1. There is the question that I had asked a few in this in my previous lecture, but now I now I (( )) another question to this. That is suppose the item selected has been found to be a defective 1. Then, what is the probability that this item has indeed being picked up from say particular box say box 1 what is a probability?

(Refer Slide Time: 08:17)

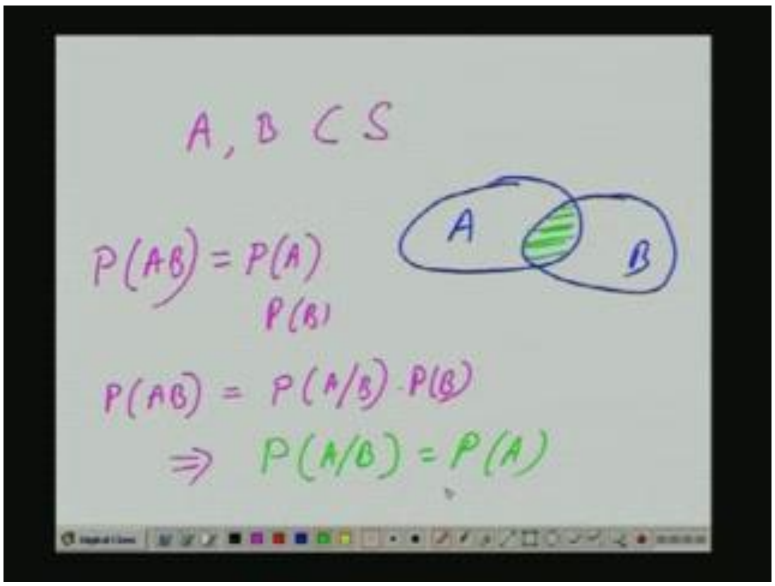



For that we will be using this base rule we know that. So, we consider box 2 what is the probability that the ball has been removed from box 2? This we know 0.4 question is what is the probability that I mean being that the ball that the item that has been removed is a defective 1 subject to that, now what is the probability that, the item has been picked up from box number 2, that is this probability. That is equal to what that is equal to by what we studied earlier first if you take the multiplication this is the joint probability. If you take this PD by B 2 into PB 2 this is a joint probability that box B 2 are selected and the item removed from it was defective, but the joint probability is nothing, but P B 2 by D into PD. Or this is equal to this side divided by PD. In fact, we have established this relation in our previous class. So, I am not speaking much on this, but roughly you can remember it like this.

That is if you take the product of that 2 again you get a joint probability between B 2 and D, because is B 2 by D probability B 2 by D. And this is the probability of D is written the other way here probability D by B 2 into probability B 2 both cases we get joint probability of D and B 2. And now this figures we know, because PD has been revaluated these were given that box was B 2 the probability of these being defective was 0.4.

Probability of B 2 is 1 by 4 as we know. So, this is point 4 into 1 by 4 divided by PD which is 0 point 1 6 2 5. I think this turns out to be point 6 1 5 all right. So, this is the answer. So, towards a end of our previous class we had just introduced you to the concept of statistical independents of events. Statistical independents of events I now try to highlight that and also discuss the various properties that associated with this motion. And I will try to give at example also to explain it further.

(Refer Slide Time: 11:18)



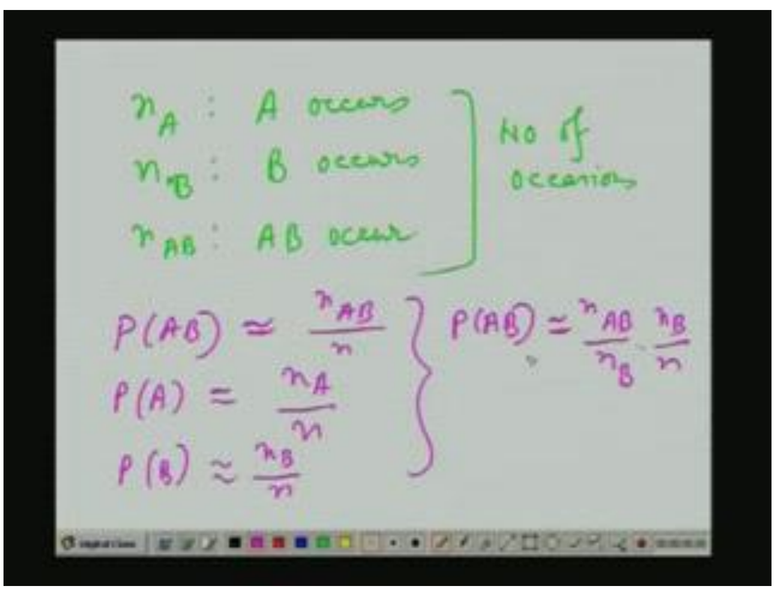

Suppose, you consider 2 events A and B, both belonging to the certain event S that is a total set S. In terms of Venn diagram you can have something like this A this is B this is the overlap area that is AB all you have studied earlier. Now, question is if A and B the events are such then if I look for the probability corresponded to there inter flexure which means there joined occurrence. Because, it is here we have both A and B occurring simultaneously both are present that is joint occurrence that is PAB if this is noting, but just a product of PA and PB.

PA is the probability of event A PB is the probability of event B, then A and B will be called statistically independent. I have tried to give you the physical meaning of this, but we hope that you can see 1 thing you can always write PAB this is always true that PA by B into PB this is always true. Irrespective to whether these 2 are statistically independent or not this we have already established this is this comes from the definition of the condition of probability in fact.

Now, if they are statistically independent the PAB is also given by these and since PB is non 0. Suppose we consider we do not count the case of PAB in 0, then obviously, PA by B is PA. In fact, how did we this come I mean this cut this possibility of PB in 0, because we are not considering any event which is a empty set. That which is impossible event only for impossible event the corresponding probability is 0 that I am not considering which means PB is a non 0 positive number less than 1 with the positive the real number.

Which means; that means, PA by B is same as PA; that means, actually that even if we first ensure that B has taken place. And subject to that try to evaluate the probability of A. The probability should be the actual probability of P. Irrespective of the fact that B has only occur. This is the meaning and this can happen in B and A are to such physical events which have no connection between them. Then whether B has occurred or not if we try to evaluate the probability of A you will always get PA. So, PA by B is PA that is the meaning of a statistically independents of A and B. Which by definition is it will like this.

(Refer Slide Time: 14:48)



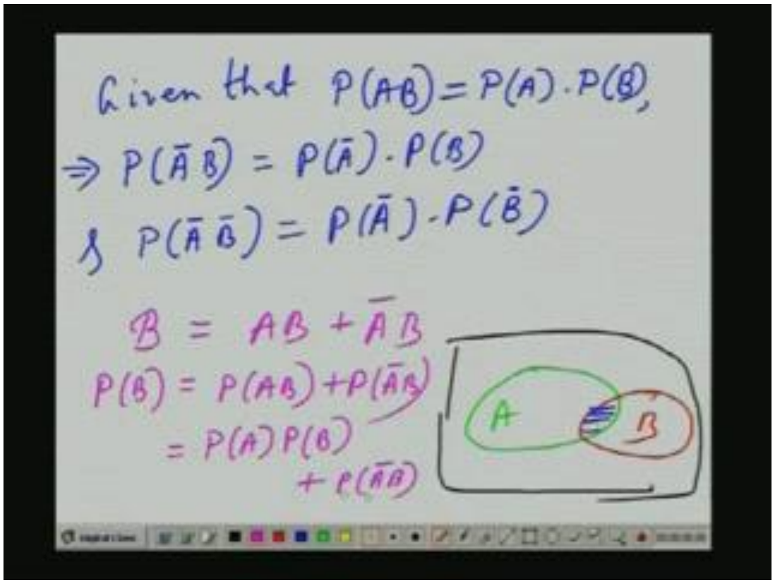
We can just look into the corresponding frequency interpretation some also. Suppose, we are conducting experiments and on nA occasions. Total number of trials is n small n on nA occasions A occurs sorry AB occur. So, number of occasions. You get case what is PAB that is simply nAB by n let me put an approximate sign, because after all this is not the exact probability, but this is just the frequency interpretation. So, nAB by n.

Now, we know what is PA that is nA by n what is PB that is nB by n and if these are statistically independent, then it would be in it would be in this PAB. We can only just write it like this nB then nB by n we can only set like this is. And what is nB by n that is PB and what is this PA by B, but this transfer to be nA by n. That means the fact that on nB occasions B has occurred has got no impact on the probability of A's occurrence. It is just repetition of what I stated using symbols earlier.

So, whether B has occurred or not that is nAB by nB provided always number of trails are large nB is large. NAB by nB will turn out to be what we get as nA by nonly provided B has got no connection with A. That is B and A are physically and statistically independent other this is PB and this is PA by B, but by definition PAB. If they are statistically independents equal to PA to PB which means, this is PA and if this is PA; that means, nA by n ratio will turn out to be same as this.

That if we just takes n number of trails and find out total number of times AB has occurred and take this ratio whatever I will get. I will get the same thing if I take nB number of trails that is trails were first it is ensured that B has taken place. Then find out how many times A has taken place and take the ratio. I will still get the same value as I get here provided. Of course, these are large enough. That means the fact that B has occurred has got really no impact I mean whether B has occurred or not. I get the same probability of A then only I see that B and A are statistically independent.

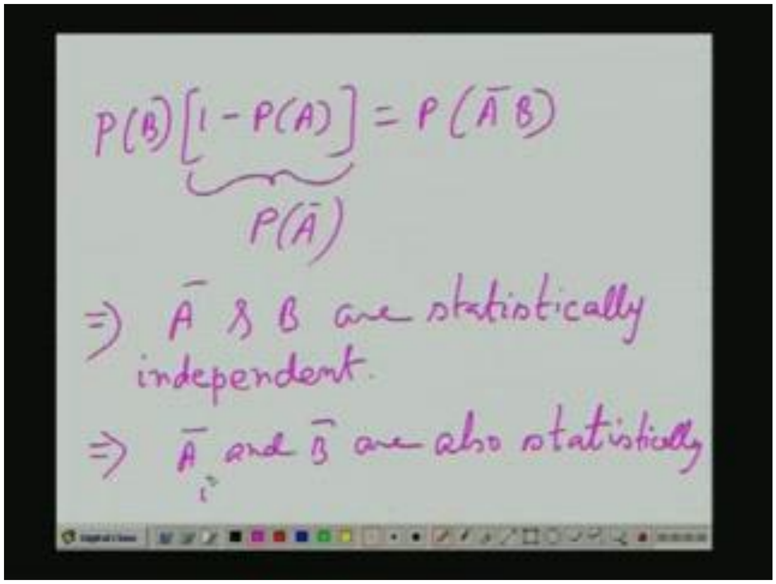
(Refer Slide Time: 19:02)




Then certain properties follow from this; given that, that is given that where 2 events A and B are statistically independent. We should have PA bar B is also written similarly that is the 2 events A bar that is A complimented B. They are also statistically independent further PA bar B bar; these 2 also are statistically independent. It is not difficult to show this. First consider this we know that any event B can be written as a union of 2 mutually exclusive sets 1 is AB another is A bar B.

If you have forgotten I just draw a Venn diagram here quickly this is A this is B. So, this is a diagram this part is AB this part is AB and this remaining portion of B is nothing, but intersection between A compliment that is the area outside A and B. Where only this portion will be absorbed right and these 2 portions are mutually exclusive. So, B can be written like this we have seen it already earlier, but since they are mutually exclusive PB is nothing but … You have given me the fact that A and B are statistically independent which means I can write like this.

(Refer Slide Time: 22:26)

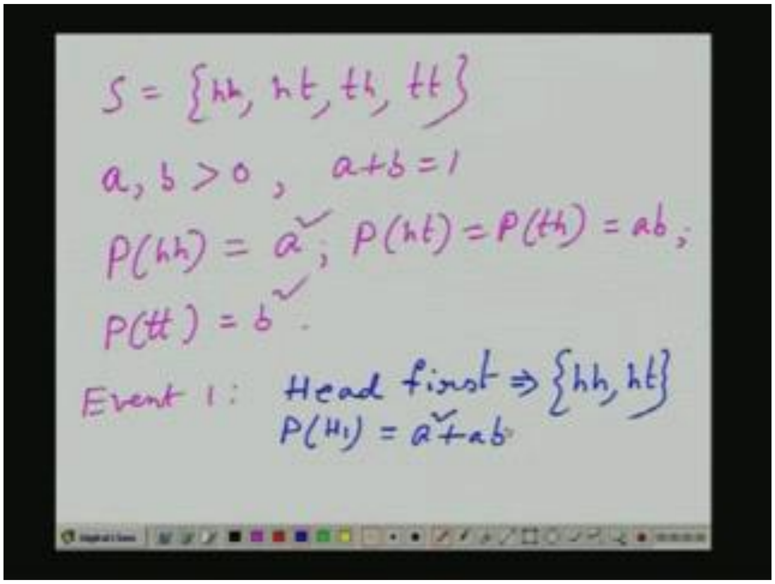



So, I take this PB to the left hand side. Take the test from 1. So, you get now what is 1 minus PA is nothing but, this is nothing, but P of A bar this we remember P of A. And P of A bar the 2 probabilities when added their value is 1, because A and A bar when put under union they correspond to the total set there is the certain event. So, it is probability is 1 which means PB which means P of A bar B is same as PA bar into PB all right this implies A bar and B are statistically.

If this logic you can apply again it is giving suppose that AB and are B A bar B they are statistically independent which means, A bar and B compliment that is B bar. They also statistically independent by the same argument there is 1 event early to us A and B and then compliment came over A P left was left as it is. And now you are given this fact that A bar and B are statistically independent.

Then keep A bar as it is compliment B. So, you get A bar and B bar by the same development again I can say that A bar and B bar also are statistically independent then repetition of the same argument. Which means, just to give you an example (( )) experiment where, we are tossing a coin just twice in successor. So, when the head comes I say the outcome H when tail comes it is T. So, there are 4 possibilities, is it not?

(Refer Slide Time: 25:10)

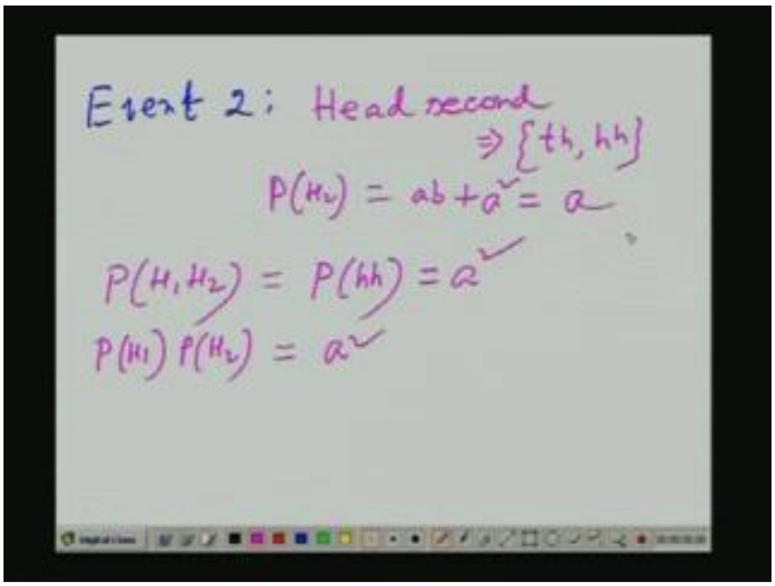



S consist of here hh that is head first head second ht that is head first tail second then th and tt. Now suppose you select 2 numbers a and b ab greater than 0 and a plus b 1. So, that they qualify for some probability actually the real numbers that is given like this. Suppose, we select any such pair in and then it is given as beforehand that P of hh is nothing, but a square. P of ht is same as p of th that is equal to ab and p of tt means b square. Then there is no conflict, because what is the total event either hh or ht or th or tt.

So, probability of the total event is what this plus this plus this plus this a square plus ab plus ab plus b square which is, a plus b whole square and a plus b where given 1 that is 1. So, that is satisfied in each probabilities having value greater than 0 and less than 1 there is no probability. Suppose this is given to us. Then we consider 2 events in 1 event head first that is I consider those subsets. These are outcomes mind you; the certain event has this possible outcomes head head head tail tail head tail tail. And then I can construct events by taking their forming the subsets.

Suppose, I now consider those case those cases where head is first, so I consider those subsets that subsets where H comes first;, so event 1 head first meaning hh ht. What is the probability of event 1? Probability I can call it say Ph 1 h 1 stands for say event 1. That will be what since they are mutually exclusive I mean these 2 and probability of this total event will be what. See either this can happen or this can happen they cannot happen together they are exclusive. So, what is the total probability total probability will be the probably sum of probability for this and this which means a square plus ab. Take a common a into within bracket a plus b and a plus b is 1 which means this is equal to a all right. So, head first its probability is a.

(Refer Slide Time: 29:23)




Similarly, we now consider second event that is event 2. Head second meaning th or again hh head second. So, here probability of this event is Ph 2 again ab plus a square take a common and you get a, because a plus b is 1 all right. Question is are these 2 events statistically independent from a physical sets yes, because whether head first or head second. These events you they really physically have got the influence each other isn't it. But mathematically does it satisfy the (( )) for statistically independent that you can verify.

So, just you check whether H 1 and H 2 are statistically independent or not what we do we take the probability of the joint event H 1 H 2. Now, what is the joint event here H 1 had hh and ht and H 2 is th and hh. So, they can jointly occur only in the form of hh if both have to occur jointly. Then only hh can take place, so that means this will be nothing, but the probability for hh, which is equal to a square, but what is PH 1 PH 2 that is again a square we have found out that in both cases the probability is a.

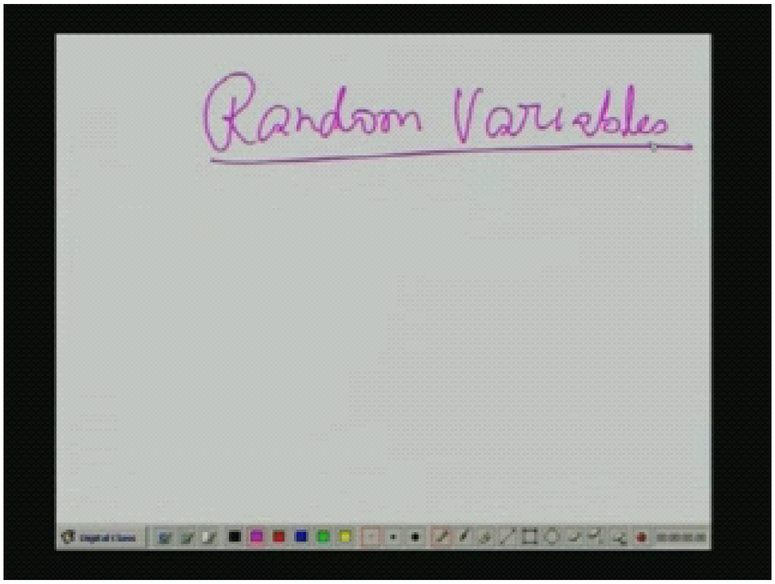
This means, this 2 events are statistically independent so far for statistical independents. So, I conclude this particular section and I will go to the following topic now, but then I think is proper that we some of rest what we did what we have done so far. We started in our first lecture with a discussion for the course we try to develop the need for, you know formal treatment to probability which we did by forming some probability (( )).

So, you towards that we first started with the concepts of set theory, because the theory of probability has been built around set theory on a set of events. Then we established we proposed or state it rather certain (( )) of probability basic (( )). Then few assumes we did satisfied by the probability and then we discuss the corollaries that come from those exhumes. We took up some examples, then we moved over to something called conditional probability then we discussed total probability or I mean base theorem.

Then we took up some examples to explain that and then we came to this notion of Statistical independents. This statistically independence is a very useful very useful concept and very widely used concept which we will study late, which we will see later as this course develops;, so this for this general definition of definition of probability.

Now, I move to the concept of random variable all right even though we say its variable actually random variable is more a function than a variable. But only problem is in our day today business I mean we tend to lose sight of its role as a function and we try to take it as a variable.

(Refer Slide Time: 34:10)




So, to give it will be formally right. Before that, let us consider this idea of a function what is a function. Function can be continuous or discrete, but always the any function it takes basically, some member of a particular set and maps it to another member of a set. A source set on which it works it is called the domain and the set which where it maps actually the domain it is called the range right.

For instance we often write a function like this x of t now here t could be t corresponds to particular set is a member of a particular set it. So, happens that often this t is a member of a set of continuous. I mean t is continuous value is parameter means that, t is member of set of continuous valued real numbers actually. This is a real line and for H element on the real line for Ht xt gives us another real number.

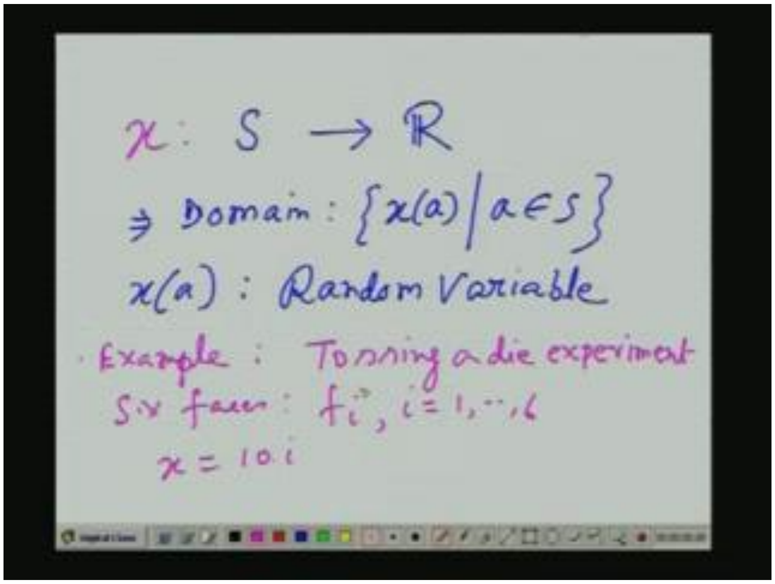
So, that is a double mapping, but it could have been a discrete it could be a discrete parameter that is t can correspond come from a set may be finite may be infinite or some discrete elements. X of t is nothing but is a function which works on those elements and gives you some value. Value could be complex or real value or any other I mean if if not real complex it could be binary also it can belong to any field. That is the general notion of function.

Now, suppose in the context of our probability I say things like this that suppose we are considering an experiment. So, experiment has got various outcomes. You consider the set of all possible outcomes that is the certain event s. Now, what I do with each outcome I assign some particular value I say that if this outcome has taken place. Let the value be this much or if that outcome has taken place let the value be that much, but the values are fixed once for all.

That means, if you go on conducting the experiment you keep getting new and new experimental outcomes. Of course, all experimental outcomes belong to that total set is, but each time you find you look at the value you see that the value is that is taking place is random. So, if you have a value chart you will see various values are coming again values are pre assigned. It is not that they are taking arbitrary values, because you already have a set of pre assigned.

Values and whatever value are now getting the correspond to that set only, but you are getting this a chart giving this random occurrence of these values. Here, we say I mean this process of assigning value to each experiment you can again view it in the same way as we have the notion of functions.

(Refer Slide Time: 37:44)



That is you can see that let x. It maps the set S to a set of numbers for making our life simple we are considering real values, but it would be complex. So, S is mapped to R. Physically, it means x is a function which works on each element of S means each experimental outcome of S for each of this experimental outcome is function picks up a value from R. R is the set of all real numbers. That means what is the domain here.

That is a is particular experimental outcomes I am considering all experimental outcomes belonging to S, where each outcome x of a is the function it picks up the particular outcome and gives you value from R. And the set of all values that forms the domain. So, x just picks up these outcomes and maps a particular outcome to particular value in the domain. So, this x of a ideally this is called a random variable. So, actually it is a function actually it is a function.

So, x of a it means basically the function as I said, but you know often what happens. We take x as some kind of variable in the sense that. X equal to what? X equal to refer this and we are measuring we are do some experiment and we are measuring something what is the x. Next time what is the x and like. So, x is treated most of an impact is a some kind of variable, because we directly look into the value of x. And that is from our point of view now from our consideration this fact there is an experimental outcome on which x is working and giving you the particular value that aspect is hidden.

For instance, suppose, there is furnace every 5 minutes we are measuring temperature of say particular section. There are various possible temperatures may be a continuous range of temperatures with certain discrete upper approximations and discrete range. There from a set of experiments when you measure is a thermometer the value that you are getting that actually is a function.

There is the this temperature you know it takes 1 of those events and this measurement rather measurement of temperature it takes 1 of the I mean states. Suppose, furnace has certain states whether state 1 or state 2 or state 3 or state 4 each states is characterized by particular temperature. And this states from the events the process of measuring the temperature (( )) by a thermometer or whatever.

What it does it picks up 1 of those states that is 1 of the events and then gives you a particular value that value is temperature T. Unfortunately in practice we will be taking T only and we will be taking T as some kind of variable, that we measure temperature now what is T T is so much degree centigrade. Next time what is T T is this much and like that.

So, for us T will appear to be some kind of variable, but actually and we in loosely call them random variable, but actually with each random variable. There is the interpretation of an experiment that goes on and the random variable is nothing, but a mapping of particular event to the particular experimental outcome rather. Not event experimental outcome to real number and passing real valued to random variables;, so a particular real number.

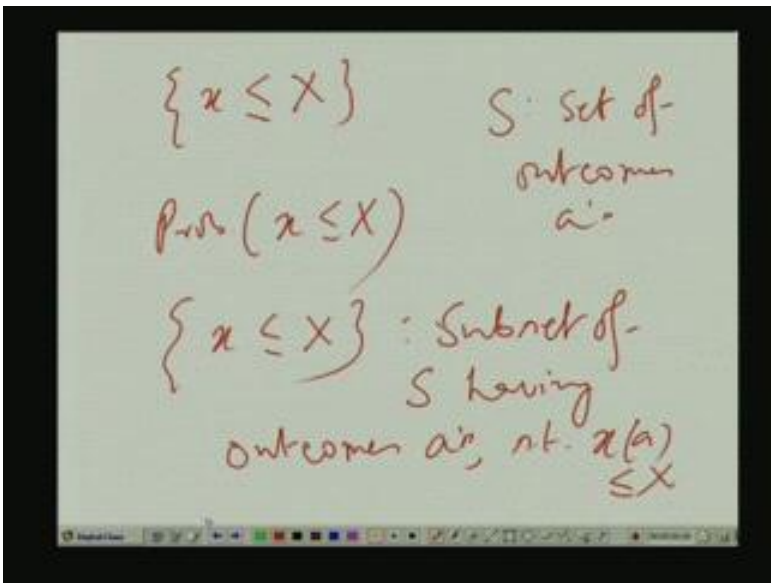
Then just to give an example, suppose you are a kind given example of the furnace, but let me give another example to be this point clear. Suppose, I take that crossing a die experiment or die has got 6 faces you the get the faces be f 1 face 1, f 2, f 3, f 4, f 5, f 6. F 1 means the top face is marked with 1, f 2 means that top face is marked with 2 and likewise. And each of the states each of the faces has probability 1 by 6, because it is an unbiased die all right.

What we do we conduct an experiment 1 of the faces show up and based on what we hav. Suppose, we generate a particular value and we generate by some mechanism in this case the mechanism is that is let me write it out. We have got 6 faces fi I from 1 to 6. So, any of the fi will be each of the each fi is a particular experimental outcome. So, there are 6 outcomes and total set is nothing, but a set consisting of f 1 f 2 f 3 or f 6.

What we do the moment the particular event takes place particular outcome takes place say fi we assign we find out this variable x is equal to 10 i. Which means when f 1 takes place the particular value that will be 10 if f 2 takes place it will be 20 then 30 40 50 60. So, every time we will conduct this experiment you get 1 particular value from that set 10, 20, 30, 40, 50, 60. So, in this case your function it maps each outcome there is a function, which maps each of the outcome to a particular number f 1 to 10 f 2 to 20 like that.

That function is actually described like this actually I should have written x as a function of I, because i is the event. But as I told you often we really do not I mean we just deal with this particular variable and often that mechanism of experiment and particular experimental outcomes and all (( )) that remains hidden from us. That is why we most often I mean treat the variable as x itself where that calling it at x of some event x of some experimental outcome I.

(Refer Slide Time: 46:31)




Then comes this issue on particular notation certain notations are important here. Now, what does it mean? We know x is a random variable the x. Now what is what is x actually? That we know that there is set S it consists of all experimental outcomes set of outcomes a yes. Now, in each outcome as I told you we associate I mean we have a function actually x should be strictly a function of this outcome. That is on each outcome there is a function which works and gives you a number.

So, actually it is a dot x it should be x of a, but loosely as I told you earlier we omit this dependents on a we simply say x. In any case x says stands for a number which is associate with the particular outcome for each outcome the corresponding x gives you 1 particular number. And since, these outcomes are occurring randomly that is every time you perform an experiment you get some outcome or another.

The random the value that x takes is also done, because of you at every trail or at a re experiment you get some new value or different value. So that values occur also the numerical values also occur randomly that is why it is called random variable. Now, coming to this what does it mean? X is a given number. And what I say X less than equal to X and I put parenthesis what does it mean now, this notation we have to first make clear, but before that let us tell you this.

That it probability theory we often try to ask this question that. Given a number X what is the probability that small x the random variable is less than equal to X. As you know we have defined probability not on numbers, not on x or X, but we have defined probability over events. There is a set of all possible outcomes then subset of that set from the events with each event we have a probability and we have given the (( )) depends on probability already.

So, events probabilities always associate associated with the events not with the numbers, but in probability theory or random variable theory. Since, we talk in terms of random variable where x is actually of a function. But we do not I mean do not show that functional form, but I mean just write X, but actually it means that there is some experimental outcome upon which I mean based that outcome x takes some value and in that sense X is random.

So, in that background question is what is the meaning of this or rather if I want to say that what is the probability? What is the probability that random variable x is less than equal to capital X how to define it. You understand my point my point is. So, far probability the notion of probability has been defined on set that is events not on numbers.

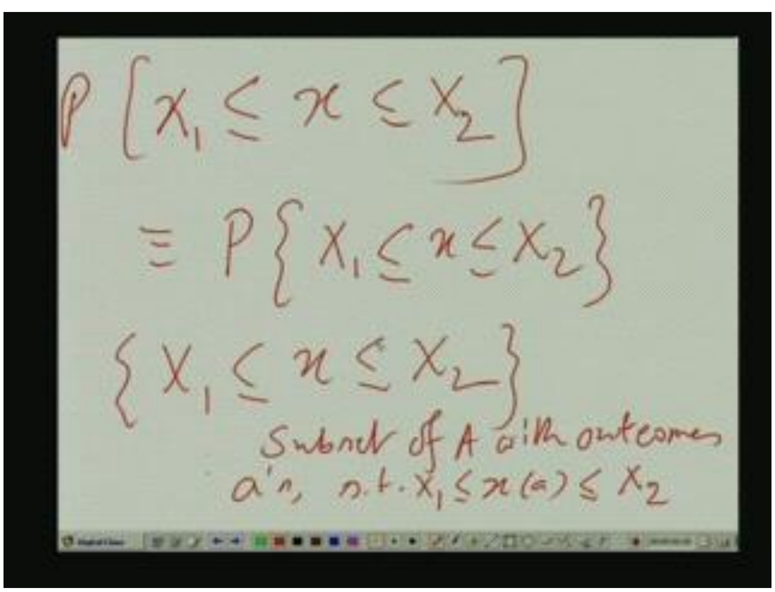
So; that means, if you have a situation like this condition that given to us like this the x is the random variable. But, I mean it takes values less than equal to a given number X this situation somehow we have to convert into the language of I mean into I mean equivalently into this events. That is we have to somehow find the interpretation of this a situation in terms of the events that is sets or subsets.

Then we can talk into in terms of probability, because probability is not defined on these numbers this is not defined on the variable X, so far it is define only on sets. So when you write this x less than equal to X we actually mean an event and when I say probability of small x less than equal to X. It actually means probability of that particular event which is associate with this. Now, what is that event that event means as you know the set consist of several outcomes, with each outcome I have got value for random variable X.

Now, from all those outcomes now isolate only those for which the value that X takes is either equal to X X or less than X. And collect them see that forms a subset of S which means it is an event, because here is a subset consisting of some outcomes. So; that means, this is nothing, but a subset of S having outcomes a or s. So, that x of a actually as I told you random variable x its actually a function it takes that outcome and gives you a value see x of a is less than equal to X. So, I find out those outcomes only for which x takes values less than or equal to X.

Set of those outcomes definitely is subset of face at differently is an event that has the probability. So, probability of small x less than equal to capital X equally that probability of that event. Which event that consists of those outcome for which x takes values less than equal to capital X.

(Refer Slide Time: 52:23)



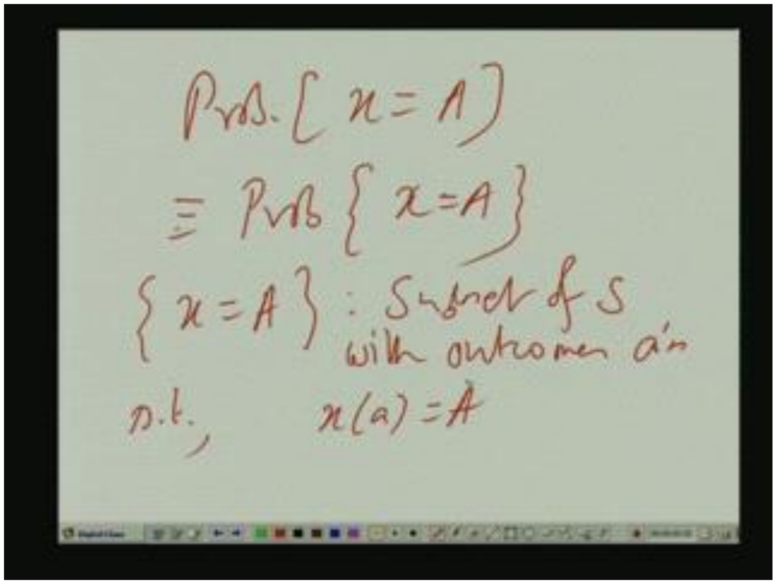
Then that sometimes we ask this question. Given 2 numbers X 1 and X 2 what is the probability of this what is the probability P for probability what is the probability of X lying within this less than equal to X 2 greater than equal to X 1. Again this statement is nothing X less than equal to X 2 greater than equal to X 1. It is purely in terms of number and variable 2 numbers X 1 X 2 and x is a random variable. That (( )) I mean cannot apply the probability on that, but I have to convert this condition into a condition involving events.

I have to convert this condition into some event and on that event definitely we have got the notion of probability working. Now, this will imply the this is same as the probability of that event just a minute. This is equivalent to probability of this event what is this event I will tell now, when I put curly bracket this becomes an event now this is what is this event this actually denotes an event. What is that event it is set of rather I would say subset with outcomes is. So, that the random variable is less than equal to it takes the value that is less than equal to X 2 greater than equal to X 1. That is from the given set S you isolate you find out only those outcomes for which X takes values within this range X 1 to X 2. Because with each outcome I have got some value associated as expressed by X which is actually a function of that is outcome.

So, I am just (( )) collecting some outcomes for which all those outcomes mind you. I am not leaving any I am collecting all those outcomes or all those elements of the set S. So, that for each element that is for each outcome if you for each of those outcomes. If you take that value of the random variable you get a value within the range X 1 to X 2, but that is again set of some outcomes that is set of that is subset of a, which is an event. And on an event we have definition of probability working.

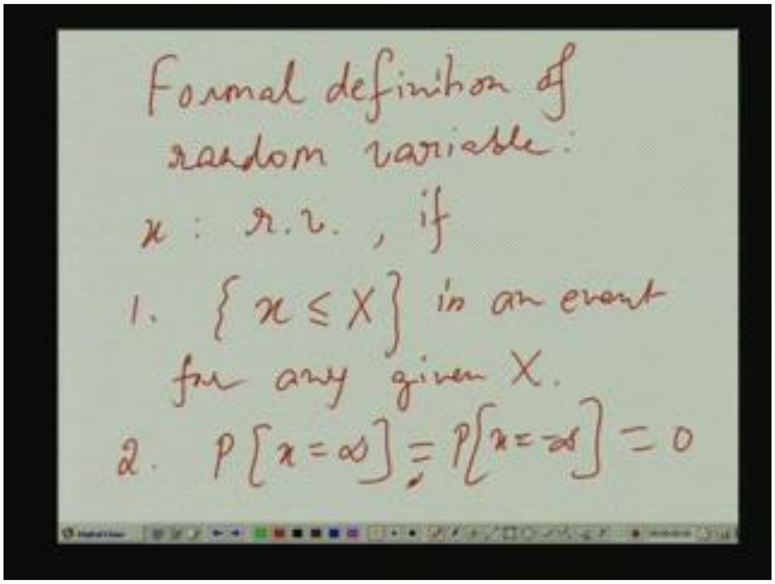
So, probability of X lying within this range capital X 2 and X 1 is same as probability of this event that is I mean probability of this event occurring. That is I mean probability of occurring probability that some outcome occurs. So, that X of that outcome there is a random variable takes values within X 1 to X 2, that is the probability of that event.

(Refer Slide Time: 55:47)




Similarly, if I say probability of say x equal to A some number A. It means it is equivalent to probability of this event again curly bracket curly bracket stands for event here. X equal to A, but this event means x equal to is A condition, but when I put within curly bracket it means an event. Event means some sort of subset S it it gives rise to a subset of S or it indicates the subset of S what subset that is such a subset of S with outcomes S. So, that x a is equal to A.

(Refer Slide Time: 56:45min)




Finally formally we define the random variable like this. X is a random variable or rv rv stands for random variable if number 1. This event this is an event for any important is any given X. Means you give me number X and then you find out x less than equal to X within curly bracket means, you try to look for some outcomes of S. Or those outcomes of S for which the random variable takes the value less than equal to capital X. Less than equal to capital X.

Clearly, that is I mean this portion should be an event that we should really be able to find out that outcomes. So, this must be an event. Number 2 is the 2 things probability of x taking value infinity and probability that should be equal to probability of taking the x value minus infinity both should be equal to 0. That is the formal definition of random variable that is all for today.