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

Module No. #01
Lecture No. #07
Random Variables

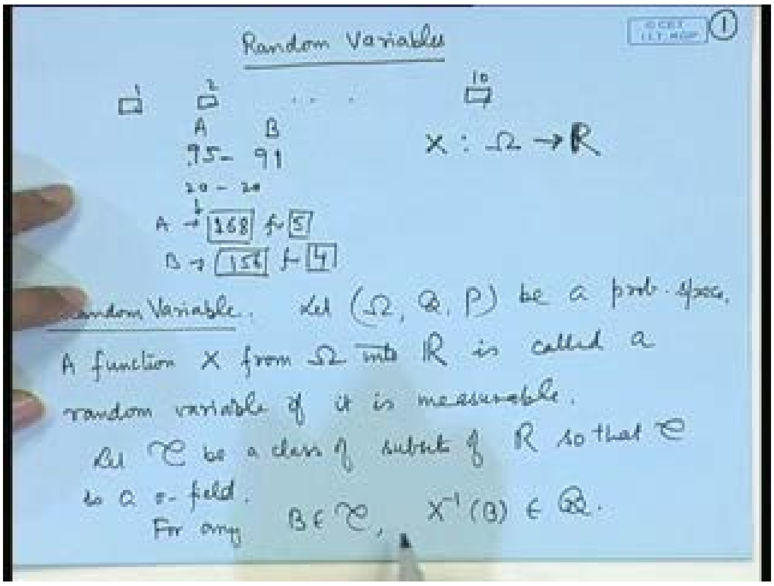
So, far we were discussing the laws of probability so, in the laws of the probability we have a random experiment, as a consequence of that we have a sample space, we consider a subset of the, we consider a class of subsets of the sample space which we call our event space or the events and then we define a probability function on that.

Now, we consider various types of problems for example, calculating the probability of occurrence of a certain number in throwing of a die, probability of occurrence of certain card in a drain probability of various kinds of events.

However, in most of the practical situations we may not be interested in the full physical description of the sample space or the events; rather we may be interested in certain numerical characteristic of the event, consider suppose I have ten instruments and they are operating for a certain amount of time, now after amount after working for a certain amount of time, we may like to know that, how many of them are actually working in a proper way and how many of them are not working properly.

Now, if there are ten instruments, it may happen that seven of them are working properly and three of them are not working properly, at this stage we may not be interested in knowing the positions, suppose we are saying one instrument, two instruments and so, on tenth instrument, 1 2 up to 10, we are not very particular whether instrument number one has failed or two has failed or ten has failed; that means, which three of them have failed, may not be of interest, rather we are interested in the total number of the instruments, which have failed and total number which have which are working.

(Refer Slide Time: 02:04)



Ah you look at say a game of basketball, now in a game of basketball after the completion of the play, the score line will read something like say 95, 91 a team A may score 95 and team B may score 95, the outcome of the match is that team A has beaten the team B, to A particular observer this particular phenomenon or this basketball match has the final outcome, that the team A won the game or the final scores 95 and 91. There as for a audience, they may be interested, they may be looking at which player scored how many baskets, what was the percentage of the position of the team A over the team B, over the ball etcetera.

So, similarly, if you consider a game of cricket, then suppose we are looking at a game of 20- 20, then team A scores say 170, 168 runs for say loss of 5 wickets in 20 overs, team B may score say 156 runs for the loss of say 4 wickets at the end of 20 overs, the outcome of the match is that team A won over team B or if you look at the total scores, then these are the total scores.

If you look at the total number of wickets, then these are the total number of wickets lost, we may not be interested in seeing that out of 20 overs, there are 120 balls at which ball, how many runs were scored, how many balls were dot balls.

This although these are important data, they may be useless for a certain person who is interested only in knowing, what is the total number of score, total wickets taken or a particular player scored, how many runs or how many wickets he has taken.

Therefore for example, if we are looking at say weather, now the weather is a complex phenomena so, if we look at say monsoon season, then during a full monsoon season, the amount of rain varies day to day from region to region from time to time.

That means in a particular day also a particular portion of the day may be dry, particular portion may be cloudy and particular portion may be rainy, where as for the weather scientist individual phenomena may not be of interest, he may like to know that how much total rains are there, in different zones of the country during the entire monsoon season.

If we are looking at say crop of a certain variety and we look at the total yield say wheat, then we may be interested in knowing that, how much total crop of wheat was produced in India during 2008 to 2009, we may not be interested in knowing which plot of which farmer gave which, how much yield or how many facilities were there because of which something happened.

The total amount may be of interest to us, in effect what I am trying to convey is that, in order to effectively study a certain random phenomena or a certain random experiment, it is require to quantify the phenomena.

That means for every outcome, we associate a number a real number and then, we can study the probability distribution of that that brings us to the concept of random variables.

So, what is a random variable then roughly speaking, random variable if we give a notation X, then it is a function from omega into R.

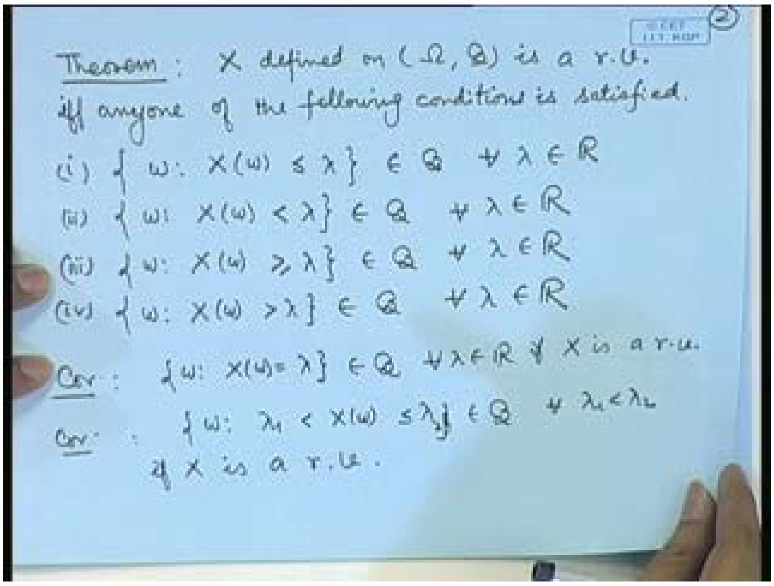
That means for every outcome we associate a real number for example, runs scored by player the baskets hit by A, a by A player in a basketball game, the yield of the certain crop, the amount of rainfall in a certain region etcetera so, all of this will correspond to certain event in the sample space and we are associating certain real number there.

So, in particular any random variable is a real value realization of a sample space so, it is a function from omega into R; however, to be more precise, because we have defined the probability function over a sigma field of subsets of omega.

So, it is important that when we want to talk about the probabilities of associated with those numbers, then the inverse images of those subsets must be in the set B so, this is called the condition of measurability and formal definition will then become the formal definition random variable is so, let omega B p be a probability space. A function X from omega into R, R is a real line is called a random variable, if it is measurable.

So, to explain the condition of measurability, let C be a class of subsets of R so, that C is a sigma field, then for any B belonging to C, X inverse B must belong to script b, this is the condition for measurability.

(Refer Slide Time: 09:18)



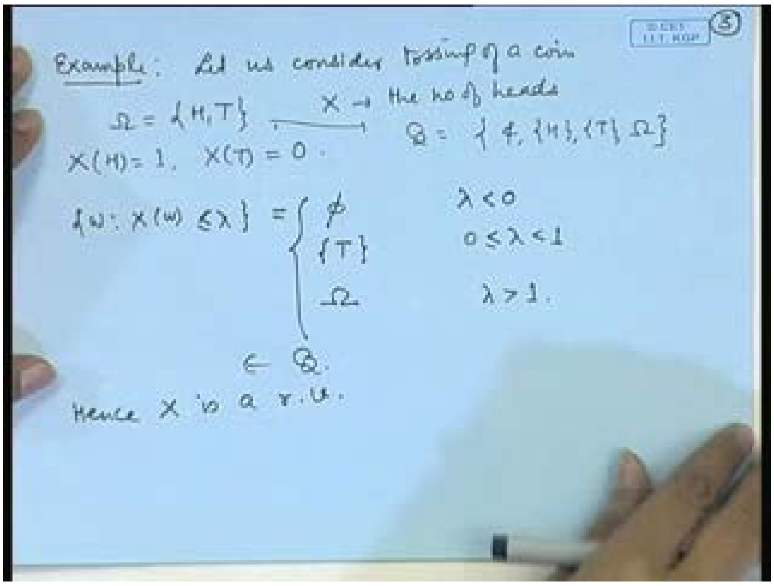
An equivalent condition here is that, we can state it in the form of a theorem so, X defined on omega B is a random variable, if and only if any one of the following conditions is satisfied, the set of omega such that X omega is less than or equal to lambda, belongs to B for all lambda, the set of all those sample points such that X omega is strictly less than lambda belongs to B.

The set of all those points such that X omega is greater than or equal to lambda belongs to b, for all lambda belonging to R, the set of all those omega such that X omega is strictly greater than lambda belongs to B

So, if any of these conditions is satisfied, then X is a measurable function and therefore, it is a random variable, as a corollary to this the set of all those omegas such that X omega is equal to lambda, then this is also for all lambda belonging to R, if X is a random variable.

Similarly, if I consider the set of all those omegas such that, if X is a random variable so, these are some natural consequences, let me just give a example here to show that, how will you check the condition of measurability.

(Refer Slide Time: 11:50)



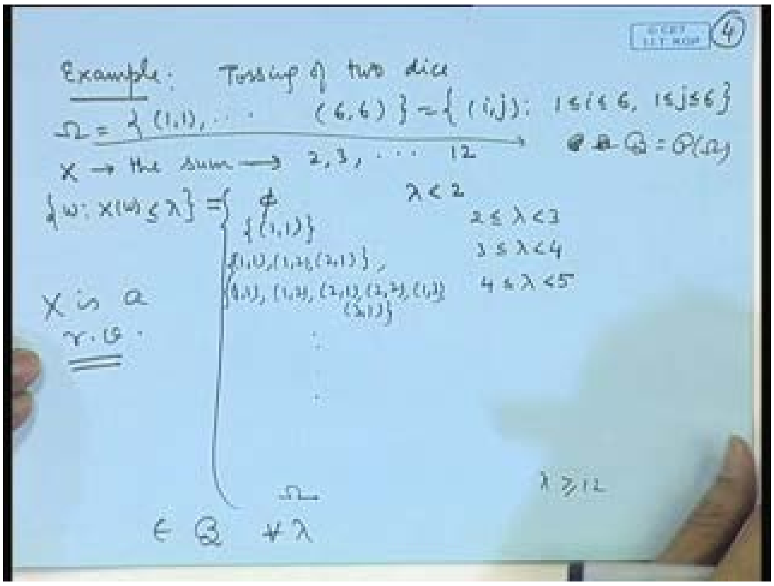
Let us consider say tossing of a coin, let us consider the random experiment of tossing of a coin so, the sample space is here H T, we define the random variable X as the number of heads so, naturally here X H is equal to 1, and X T is equal to 0 so, if we look at the set omega such that X omega is less than or equal to lambda say, now if lambda is negative, then there is no point for which X omega is satisfying that condition less than or equal to lambda.

That means this is an empty set, if I take lambda is equal to 0, then X T is equal to 0; that means, there is a point here, which is satisfying the condition that X omega is less than or equal to 0, you can observe that this condition remains true for lambda up to one; that means, strictly less than one, because the next value that random variable takes is actually equal to 1.

So, when it eventually takes one then there are two points, that is H and T which satisfy the condition that X omega is less than or equal to 1 or any point below above that so, this becomes full omega for lambda greater than 1.

So, if my sigma field here is consisting of say phi H T and omega, then you can see that this set is always belonging to B; hence this X is a random variable.

(Refer Slide Time: 14:08)

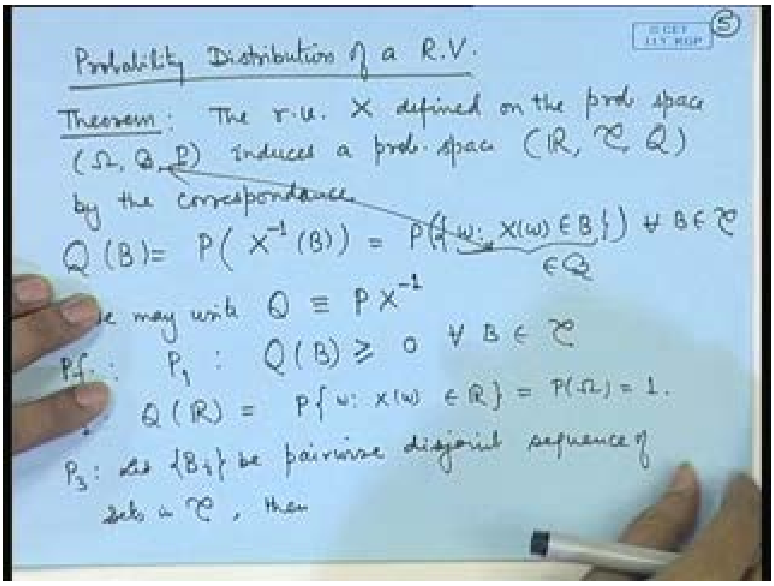


Let us take another example, say tossing of say two dice, here sample space is say 1 1 and so on up to 6 6 all the combinations of numbers, ij for 1 less than or equal to i less than or equal to 6, 1 less than or equal to j less than or equal to 6 all the combination of natural numbers between this. So, if I define the random variable X as say the sum observed and we want to write down the set omega say X omega less than or equal to lambda.

Now, if you look at this sum, sum will vary from 2 to 12 so, if I have lambda to be less than 2, then this set is going to be phi, if I take lambda to be between 2 and say 3, then I have the 0.11 here. If I have lambda between 3 and 4 greater than or equal to 3 but, strictly less than 4, then we will have 1 1 1 2 2 1, if I have lambda greater than or equal to 4 but, strictly less than 5, then this will become equal to 1 1 1 2 2 1 2 2 1 3 3 1.

Like that, if I say lambda is greater than or equal to 12 then this will be full omega so, if I consider my sigma field here, as power set of omega then this will belong to B for all lambda and therefore, X is a random variable. The purpose of considering this measurability condition is that, when I want to talk about the probabilities of these events in terms of random variable, then they should be well defined in the original sample space.

(Refer Slide Time: 16:58)

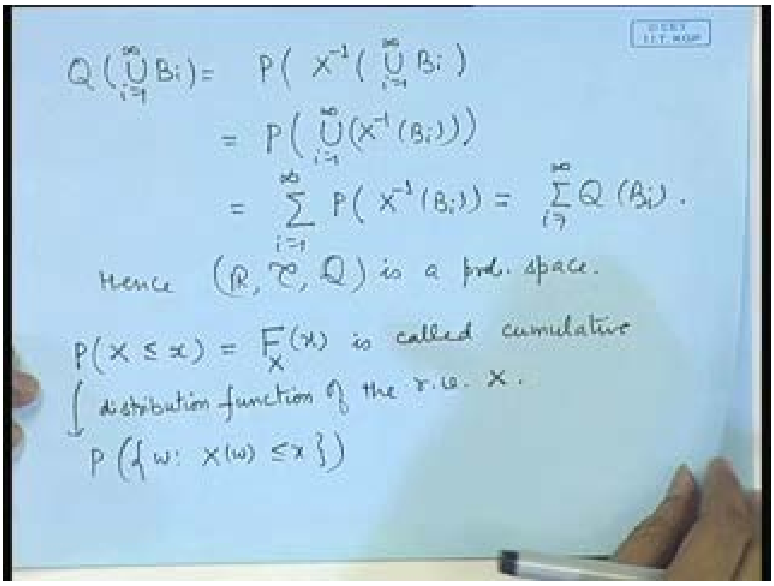


Next, we talk about the probability distribution of a random variable, we have the following result the random variable X defined on the probability space, omega B P induces a probability space, which we call R C Q by the correspondence. Q B is equal to probability of X inverse B, which is equal to probability of the that the random variables takes values in the set B, for all B belonging to C, since it is a measurable function X is a measurable function therefore, this set is belonging to script B here and therefore, this probability is well defined..

We may write Q as P X inverse, now immediately one should be interested to check whether it is a valid probability function or not; that means, it is satisfying the three axioms of probability, we claim that this is true, let us look at the first axiom so, Q B since it is equal to a probability in the original probability space, it is always greater than or equal to 0.

If I consider Q of R then it is equal to probability of the random variable taking values in R, naturally it is probability of omega and that is equal to one, if we consider bis as pair wise disjoint sequence of sets in C then, Q of union bi I is equal to 1 to infinity is equal to probability of X inverse union of bi.

(Refer Slide Time: 20:04)

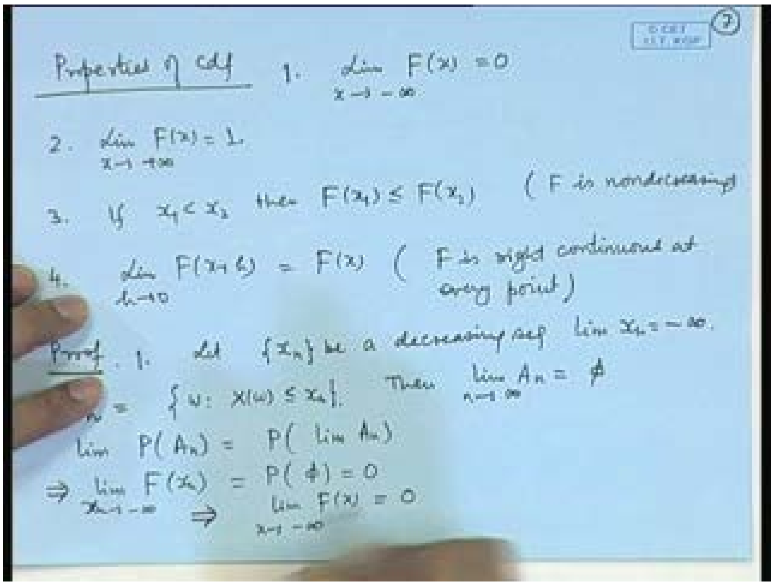


By the definition of inverse, this is if bis are disjoint then X inverse bis are disjoint and therefore, by applying the countable additivity axiom and this is exactly equal to Q of bi so, all the three axioms are satisfied and therefore, R C Q is a probability space. So, this Q is called the probability distribution of the random variable so, let us look at this example of head and tail so, here if we say that the coin is fair, then you will have probability of H is equal to half and probability of tail is equal to half.

Therefore if I say probability of X is equal to 0 that is equal to half and probability of X equal to 1 is half so, this is giving you the probability distribution of the random variable x, since Q is a set function and here I am looking at the real numbers so, there is another function which we consider continuously and it is called distribution function so, we define probability of X less than or equal to X as say F X of x, this is called cumulative distribution function of the random variable X.

This basically this is the probability of all the sample points such that X omega is less than or equal to X so, this is an abbreviation we will be writing down like this, this function is quite useful in talking about all types of probability statements of the random variable x.

(Refer Slide Time: 23:33)

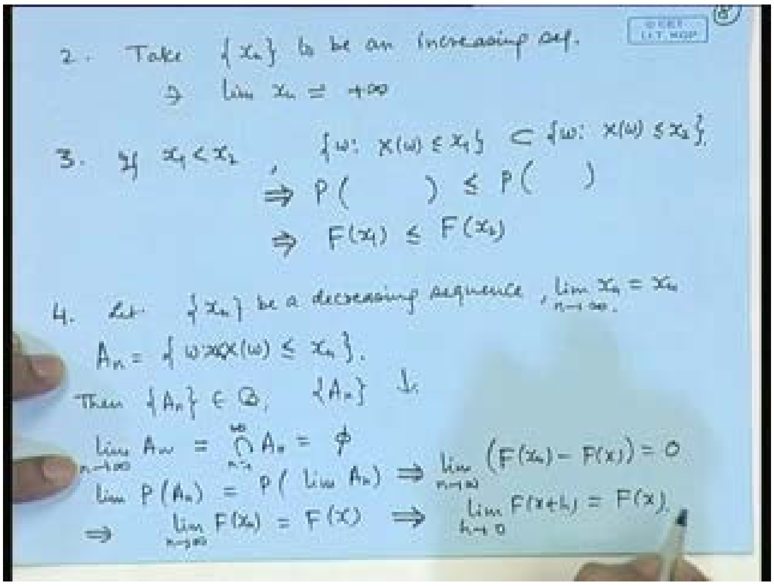


So, first of all this cumulative distribution function satisfies certain important properties, properties of cdf, the first property is that, if I consider limit of F x as X tends to minus infinity that is 0, if we look at limit of F x as x tends to plus infinity it is equal to 1, if x 1 is less than x 2 then, F of x 1 is less than or equal to F of x 2, if I consider limit of F x plus h, as h tends to 0 then it is equal to F of x, that means, continuous from the right at every point, this is F is non decreasing right continuous at every point, to look at the proves of this statement, we can consider the basic statements of the probability basic properties of the probability.

So, let us see to prove the property one, let a n be a sequence of, let x n be a decreasing sequence such that limit of x n is equal to minus infinity, if I consider the sets a n as the set of all those points such that x omega is less than or equal to x n, then limit of A n is equal to phi as n tends to infinity, because if X n is a decreasing sequence then A n will be a decreasing sequence and if I take limit of an as n tends to infinity, this will be the empty set.

So, now limit of probability of A n is equal to probability of limit of A n, so this is limit of probability of A n is nothing but, F of x n and on the right hand side, I will have probability of phi that is equal to zero, so if I am considering X n tending to minus infinity then this is equal to 0, so if along all the decreasing sequences minus infinity, limit of F x n as x n tends to minus infinity is 0, then this implies that limit of F x as x tends to minus infinity must be 0.

(Refer Slide Time: 27:05)



In a similar way we can consider a sequence, which is increasing to plus infinity and then if we take the limit, take x n to be an increasing sequence such that limit of x n is equal to plus infinity, if we take this and apply the argument given in the previous one, then here we will have limit of A n is equal to full space and therefore, limit of omega limit of A n will be omega and here you will get equal to one.

To prove the third property, if I have x 1 less than x 2 then the set omega such that, x omega less than or equal to x 1 will be a subset of the set of all those omegas such that, x omega is less than or equal to x 2 so, by monotonicity of the probability function, probability of this set will be less than or equal to probability of this set, which is equivalent to saying that F of x 1 is less than or equal to F of x 2.

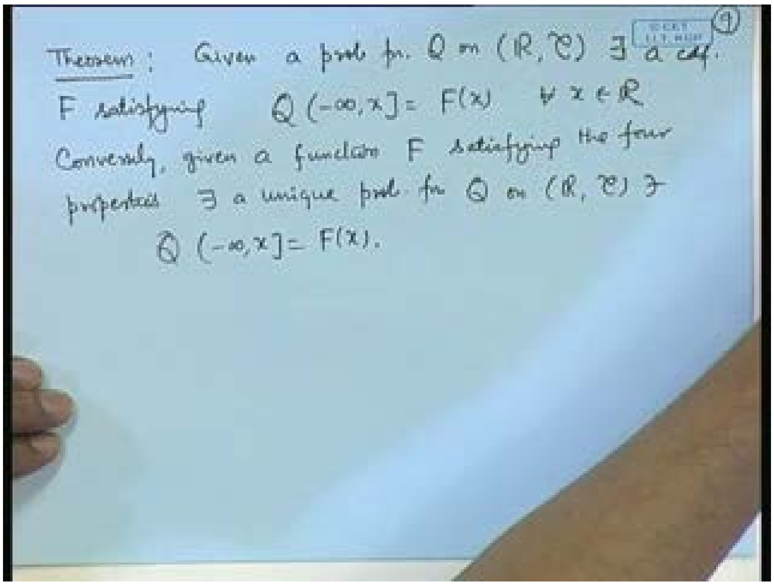
So, the cumulative distribution function is a non decreasing function, to prove the continuity from the right let us consider a sequence let x n, be a decreasing sequence with limit of x n is equal to say x as n tends to infinity, now consider the set A n as the set of all those points such that x omega is less than or equal to x n, then these are measurable sets A n is a decreasing sequence, they belong to b A ns are decreasing and if I consider, limit of A n as n tends to infinity then it is equal to, now, you look at the definition here I am considering this; however, here I should put x less than this. So, if I put this one then since this is strictly x is strictly less than x omega, so, if I take the limit here this is going to be an empty set.

That is intersection of A n as n is equal to one to infinity is an empty set, because if I consider any point x is not included in any of them therefore, it cannot be included in the intersection so, if I take limit probability of A n is equal to probability of limit of A n, then this implies that limit F of x n minus F of x as n tends to infinity is equal to 0, which means limit of F of x n, n tending to infinity is F x.

Now, here x n was any sequence which was decreasing to x that means, a sequence from a right side of x, now if it is true for any arbitrary decreasing sequence to x, then this implies that limit of F x plus h as h tends to 0 must be equal to F of x, this prove that the function x is continuous from right, this cumulative distribution function is quite an important function in order to define the distribution of a random variable.

In fact, these four properties which we have given as properties of a cdf, are also the characterizing properties of a cdf that means, if there is a function which satisfies these four properties then it will be cdf of certain random variable.

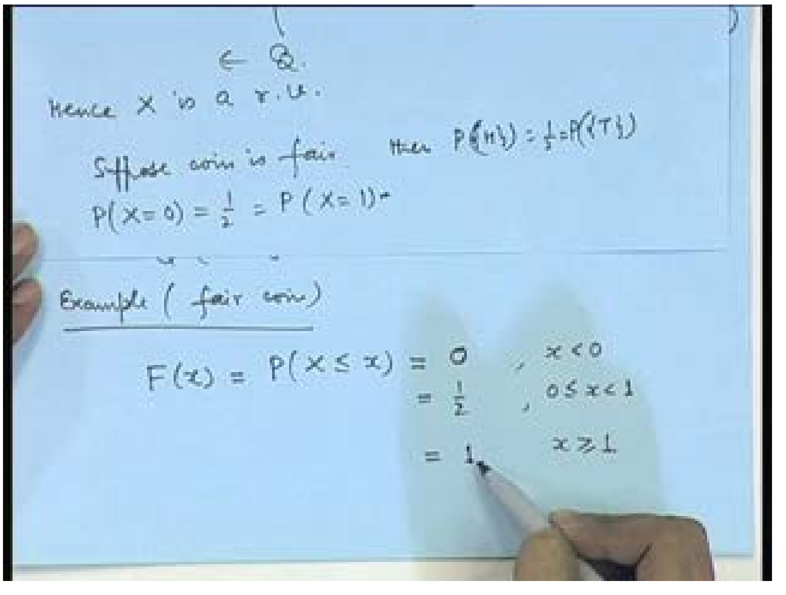
(Refer Slide Time: 31:53)



So, I state this in the form of the following theorem, given a probability function Q on say R C, there exist a cdf F which satisfies Q of minus infinity to x, is equal to F of x for all x belonging to R, conversely given a function F satisfying the four properties, there exist a unique probability function Q on R C, such that Q of minus infinity to x is equal to F of x. So, this shows that cumulative distribution function is, it has a 1 to 1 correspondence with the probability function, defined for a random variable.

And it can be actually treated as equivalent to the probability distribution of the given random variable so, let us consider certain examples here, go back to the example of head and tail and if we consider the fair coin example.

(Refer Slide Time: 34:02)

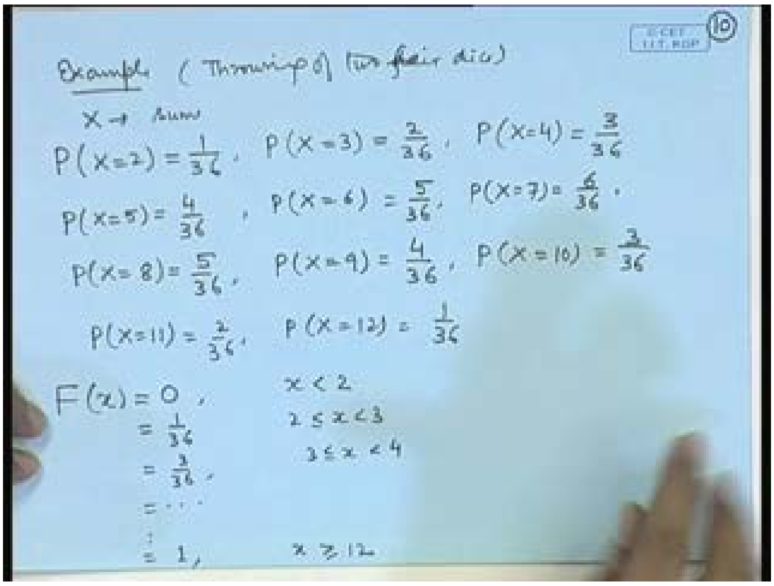


Here if we consider F x, then we are looking at the probability of X less than or equal to x so, since the random variable takes the smallest value as 0, this probability is 0 for x less than 0, the probability becomes half when x becomes 0 and it remains so, till x is strictly less than 1, it becomes 1 for x greater than or equal to 1, so, you can see here, this function is continuous on the interval minus infinity to 0 open interval, open interval 0 to 1 and open interval one to infinity.

If we look at the change points, that is at x equals to 0, the right hand limit at x equal to 0 is half, the value at x equal to 0 is half so, its continuous from right at 0, if we look at the right hand limit at 1 it is 1, the value at x is equal to 1 is 1 therefore, it is continuous from right at 1, at all other points the function is strictly continuous, the function is clearly non decreasing as x tends to minus infinity F x is going to 0 and as x tends to plus infinity it is going to 1.

So, all the conditions of the cumulative distribution function are satisfied here, let us take one more example that is of a dice throwing problem, where I defined the random variable to be the sum of the two dice.

(Refer Slide Time: 35:56)



So, in that example if we make certain probability allotment consider, throwing of two dice, 2 fair dice and the random variable X was the sum, let us look at the probability allotment here so, the probability allotment here will be probability of X is equal to 2, now this is corresponding to the event that the outcome is 1 1, so, let us look at the outcome 1 1, so, there is only 1 outcome out of 36 outcomes.

So, if it is a fair coin the probability of this will be 1 by 36, if we look at probability of X equal to 3, then there are 2 outcomes 1 2 and 2 1 which are corresponding to the sum being equal to 3, therefore the probability that X equal to 3 will become 2 by 36, if we look at X equal to 4, then we will have favorable outcomes as 2 2 1 3 and 3 1, there are 3 possible outcomes.

So, in this fashion we can write down probability X is equal to 5 as 4 by 36, the outcomes will be 1 4 4 1 2 3 3 2, probability X equal to 6 is equal to 5 by 36, probability X equal to 7 will be equal to 6 by 36 the favorable outcomes will be 1 6 6 1 2 5 5 1 and 3 4 4 3, then probability X is equal to 8 will have favorable outcomes as 2 6 6 2 3 5 5 3 and 4 4.

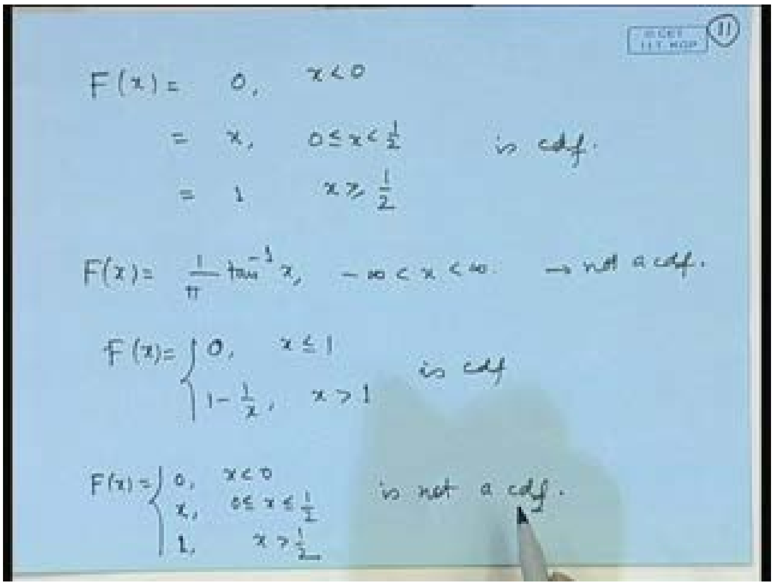
That is 5 by 36, probability X equal to 9 will have 4 favorable outcomes so, the probability will become 4 by 36, probability X equal to 10 will be 3 by 36, probability X equal to eleven will be 2 by 36, and probability X equal to 12 will be equal to one by 36 corresponding to only the outcome 6 6.

Now, if I want to write down the cumulative distribution function here, then we look at the set of values that the random variable can take it is taking values 2 3 4 5 6 7 8 9 10 11 and 12 with the positive probabilities and all others are 0 so, the starting point is x less than 2, then at equal to 2 you will have 1 by 36.

It is equal to 3 by 36 for 3 less than or equal to X less than or equal to less than 4 and so on, it is equal to finally one for X greater than or equal to twelve so, you can again see here that it is satisfying the properties of the cdf as X tends to minus infinity it is 0 as X tends to plus infinity it is 1.

It is a non decreasing function and its continuous from right, let us check the continuity from the right at the end points of the intervals, where the value of the function is changing, suppose I let back x equal to 2, then the right hand limit at x equal to 2 is 1 by 36 the value at x equal to 2 is equal to one by 36, if I take the right hand limit as x tends to 3 the value is 3 by 36.

(Refer Slide Time: 39:59)



The value at x equal to 3 is 3 by 36 so, the properties of the cumulative distribution function are satisfied here, let us take some problem say F x is equal to say 0 for x less than 0. It is equal to say x for 0 less than or equal to x less than half it is equal to one, if x is greater than or equal to half, is it a cdf, if we check the condition that x tends to minus infinity it is 0 that is satisfied.

If I take x tends to plus infinity the limit is 1, if I look at the non decreasing nature of the function that is satisfied, let us check the continuity from the right so, for the interval minus infinity to 0, for the open interval 0 to half, for the open interval half to one, it is continuous.

If I check at x is equal to 0 the right hand limit at x equal to 0 is 0 the value at x equal to 0 is 0 so, it is continuous from right at 0, let us the value x equal to half so, at the right hand limit at x equal to half is 1 and the limit as x tends to and the value at x equal to half is 1 and the right hand limit at x equal to half is also 1, therefore the function is continuous from right so, it is defining the cdf of a random variable, if we consider say F x is equal to 1 by pi tan inverse x, if we take x tending to minus infinity this goes to minus pi by 2 so, its not tending to 0 so, it is not a cdf.

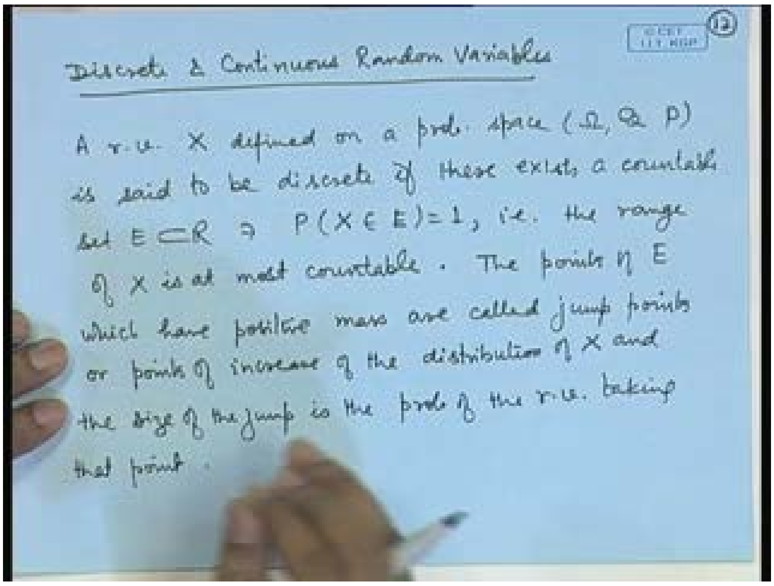
If I consider say F x is equal to say 1 0, for F x is less than or equal to 1 it is equal to 1 minus 1 by x so, R x greater than 1 once again check the conditions condition 1 2 3 they are satisfied, the limit as x tends to 1, the right hand limit is 0, the value at x equal to 1 is 0 therefore, it is a cdf this is a cdf, suppose here I change to F x is equal to 0 for x less than 0, x for 0 for x less than or equal x less than or equal to half, it is equal to 1 for x greater than half, then you see here the condition of continuity from right at the point x equal to half is violated, because as tends to right from half limit is 1 but, the value at x equal to half is half.

So, it is not continuous from right so, this is not a cdf, now we consider classification of the random variables, from the examples that we have considered for example, tossing of the coin and the random variable was defined as the number of heads so, it was taking values 0 and 1, if we consider the tossing of two dice and we took the random variable x to be the sum of faces appearing upwards, then it is taking values 2 3 up to 12.

However there may be some other random variables for example, I may be looking at the life of a bulb, then the life of a bulb could be 1 hour, it could be 1 hour 5 minutes, it could be 56 seconds, etcetera. That means it is an interval therefore, the nature of the random variable is different from the once, which we considered just now similarly, suppose x denote the number of trials to be need needed, for the first successful experiment with a certain drug to treat certain disease.

Then the number of trials may be 1 2 3 and so on that means, the set of possible values of the random variable is countably infinite so, based on this we have the broad classification of the random variables, discrete and continuous random variables.

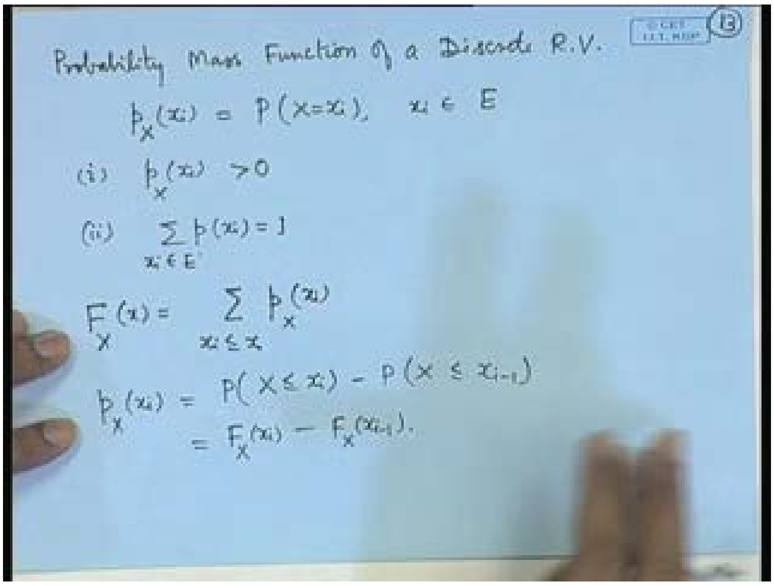
(Refer Slide Time: 44:25)



No audio from 44:24 to 44:47

So, a random variable X defined on a probability space omega B P is said to be, if there exists a countable set E such that probability of X belonging to E is 1, that is the range of X is at most countable, the points of E which have positive mass are called jump points or points of increase of the distribution of X and the size of the jump is the probability of the random variable taking that point so, the probability distribution of a discrete random variable is defined by a mass function that means, the probability of random variable taking a particular value say xi, where xi will belong to a countable set E.

(Refer Slide Time: 47:25)



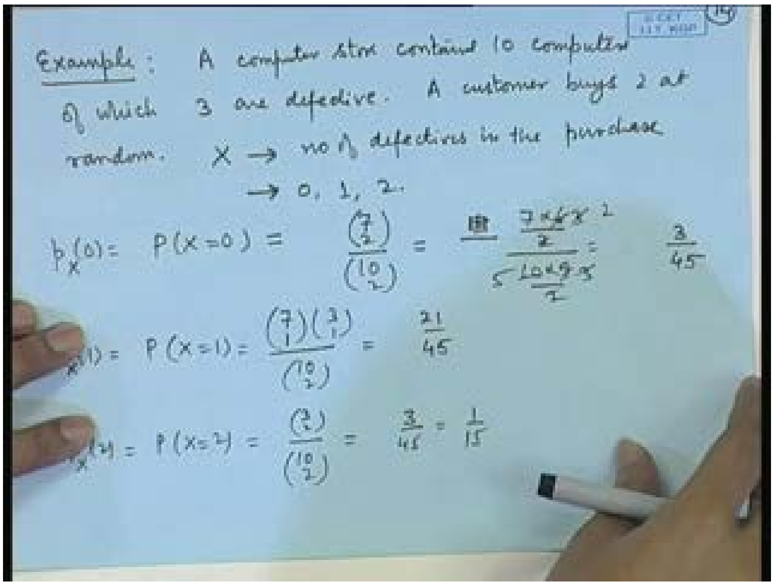
So, it will satisfy the following properties, we call it a probability mass function of a discrete random variable so, it is defined by a mass function P X xi, it will satisfy the properties so, this is actually probability of X equal to xi, for xi belonging to the countable set E so, we must have P xi strictly positive and sigma probability of xi is equal to 1, for xi belonging to so, if we look at the examples discussed earlier, the tossing of a coin the throwing of a dice etcetera.

Then the probability distribution is quite clear for example, let us look at this one, the dice throwing problem so, here it is exactly P X, you can write it in the term so, 4 3 so, you can write P X 2 is equal to 1 by 36, this is P X 3 this is P X 7 etcetera, it is clearly satisfying the properties, that P X xi is greater than 0 for xi belonging to the set here, here what is the set E, the set E is here 2 3 up to 12.

Now, from here you can establish a relation between the cumulative distribution function and the probability mass function, the cumulative distribution function can be expressed as F X is the sum of all those values of the probability mass function, for which xi is greater than or equal to x.

Conversely P X xi can be written as probability of X less than or equal to xi minus probability of X less than or equal to xi minus one so, this will become F of xi minus f of xi minus 1.

(Refer Slide Time: 50:15)



Let us take one example, a computer store contains ten computers of which say three are defective, a customer buys two at random, let X denote the number of defectives in the purchase, then the possible values of X can be 0 1 2 so, what is P X 0, now this event corresponds to that out of seven good he has chosen 2 divided by 10 C 2 so, this is equal to so, it is 7 into 6 by 2 divided by 10 into 9 by 2 which is equal to.

No audio from 51:49 to 52:05

If I consider probability X equal to 1, then it is equal to 7 C 1 3 C 1 divided by 10 C 2, if we consider P X 2 that means, both he has purchased as defective then it is equal to 3 C 2 divided by 10 C 2 so, all of these values can be calculated 3 by 45, that is equal to 1 by 15, this is equal to 21 by 45 and that is equal to 3 by 45; now here if I want to write down the probability cumulative distribution function, it is equal to 0 for X less than 0, it is equal to 3 by 45 for 0 less than or equal to X less than one it is equal to 22 by 45, 24 by 45 for 1 less than or equal to X less than 2 and it is equal to 1 for X greater than or equal to 3.

Now, you look at the jump points of this distribution so, here at X equal to 0, there is a jump of size 3 by 45 which is equal to the probability of random variable X taking value 0. I think we have made some mistake here, it is not equal to this, it is 7 into 6 by 2 that is equal to 21 and divided by 45 so, 21 by 45, 21 by 45 so, this is equal to 42 by 45 and 1.

So, the probability of X equal to 0 and X equal to 1 is 21 by 45 and the probability of X equal to 2 is equal to 3 by 45 that is 1 by 15, which is equal to the size of the jump at the points 1 2 and 0 1 and 2 respectively.

Thank you.