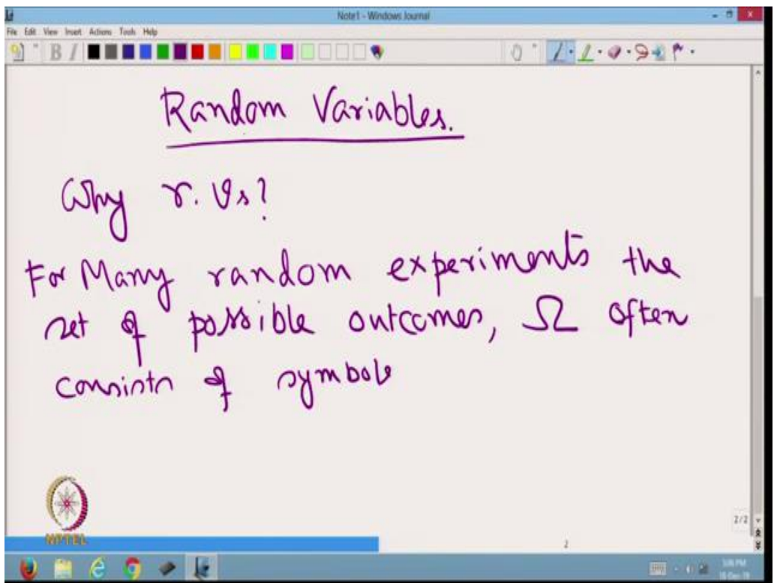
Advanced Probability Theory
Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture 6

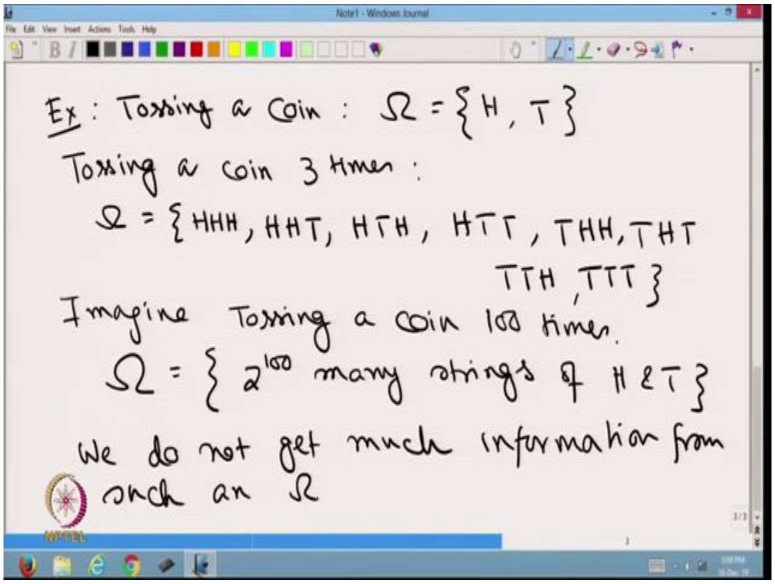
Welcome students. The MOOCs series of lectures on Advanced Probability Theory. This is lecture number 6.

(Refer Slide Time: 0:44)



As I said at the end of my last class that from today we will start what is called Random Variables. Perhaps, this is the most important concept with respect to probability. So, question comes why random variables? We have seen many random experiments for which the set of possible outcomes that is omega, the set of possible outcomes that is omega, often consists of symbols.

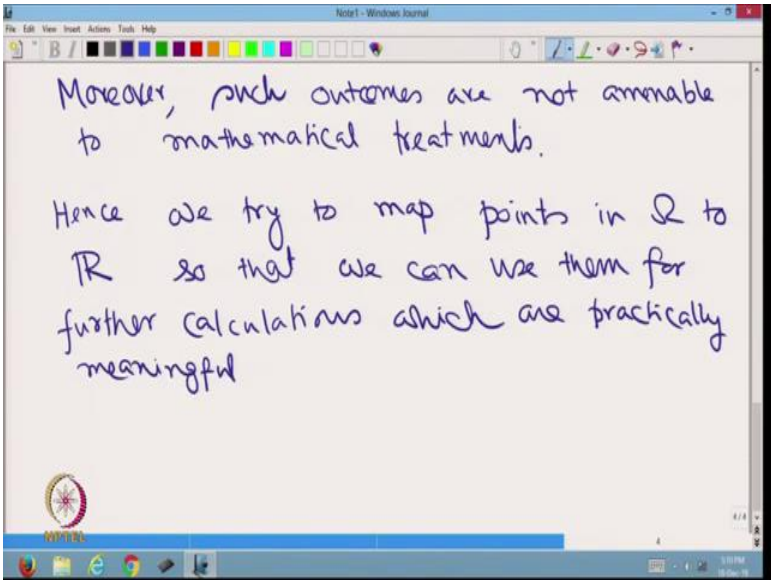
(Refer Slide Time: 1:56)



For example, tossing a coin, the omega is either head or tail. Similarly, tossing a coin 3 times will have omega is equal to HHH, HHT, HTH, HTT, THH, THT, TTH, TTT. Thus, there are 8 elements in omega, each one of them is of length 3 and each element in that string is either H or T. Imagine tossing a coin 100 times. Therefore, omega will consist of 2 to the power 100 many strings of H and T.

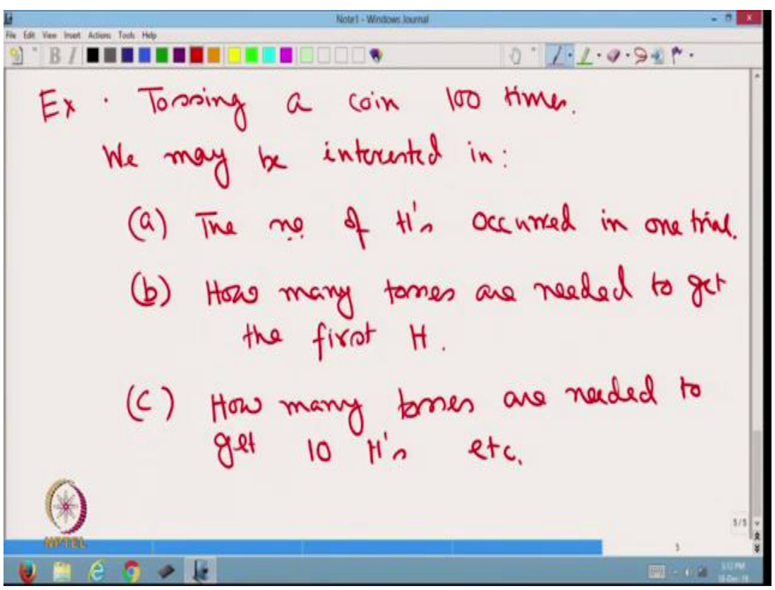
Naturally, if you have such a huge sample space, what we can infer out of it? Given any arbitrary string, nobody will be interested in which particular positions H has occurred or in which particular positions T has occurred, etcetera or in short we do not get much information from such an omega.

(Refer Slide Time: 4:42)



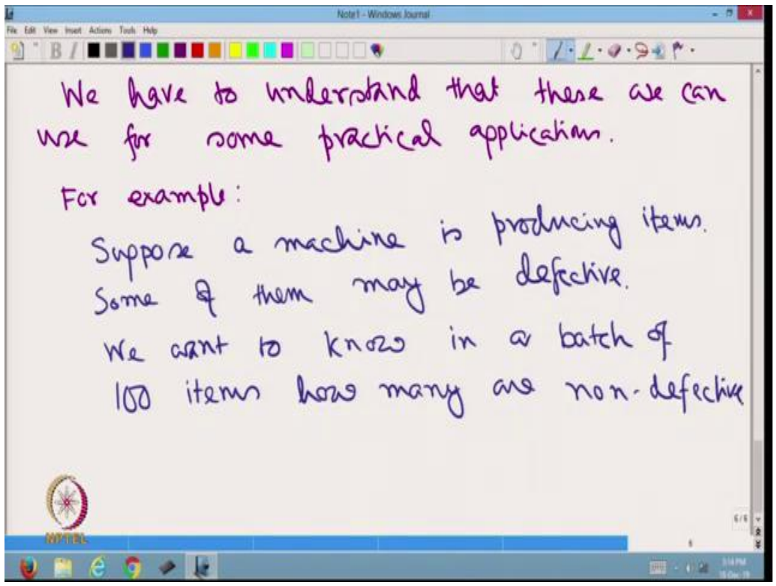
Moreover, such outcomes are not amenable to mathematical treatments. Hence, we try to map points in omega to real line so that we can use them for further calculations, which are practically meaningful.

(Refer Slide Time: 6:03)



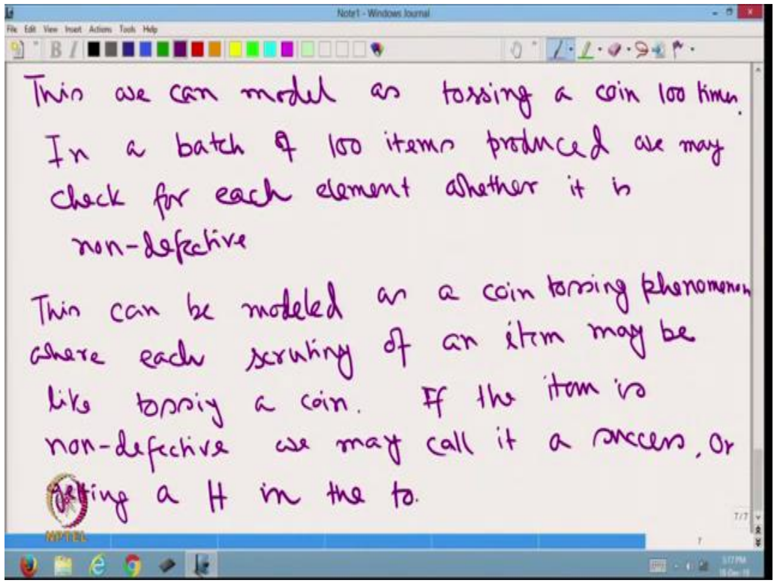
For example, tossing a coin 100 times for this random experiment, we may be interested in the number of heads occurred in one trial, note that here trial means tossing a coin 100 times and looking at its output, one may be interested in how many tosses are needed to get the first head or one may be interested in how many tosses are needed to get 10 heads etcetera? Now, you may ask me, what is the practical application of such results?

(Refer Slide Time: 7:48)



We have to understand that these we can use for some practical applications. For example, suppose a machine is producing items, some of them may be defective and we want to know in a batch of 100 items how many are non-defective. Do you not understand what I wanted to say?

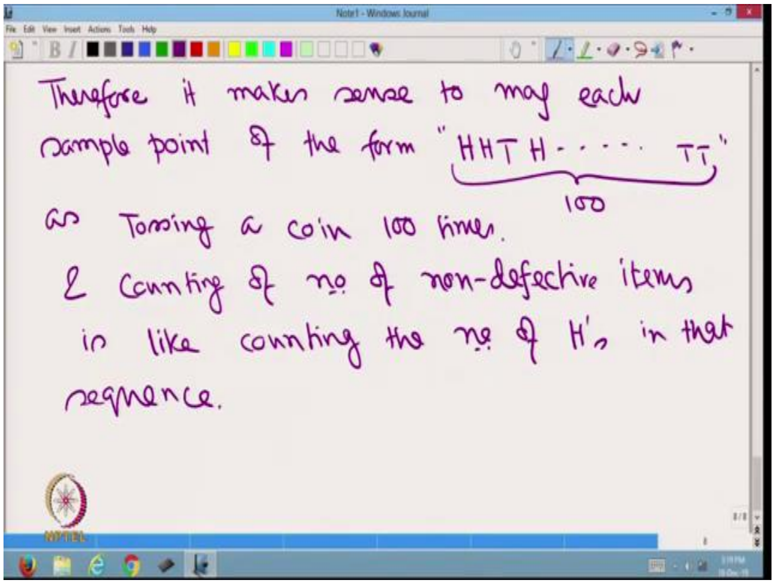
(Refer Slide Time: 9:40)



Therefore, this we can model as .tossing a coin 100 times therefore, in a batch of 100 sites produced, we may check for each element whether it is non-defective. This can be modeled as a coin tossing phenomenon where each scrutiny of an item may be like tossing a coin. And if the item is non-defective we may call it a success, or getting a head in the toss.

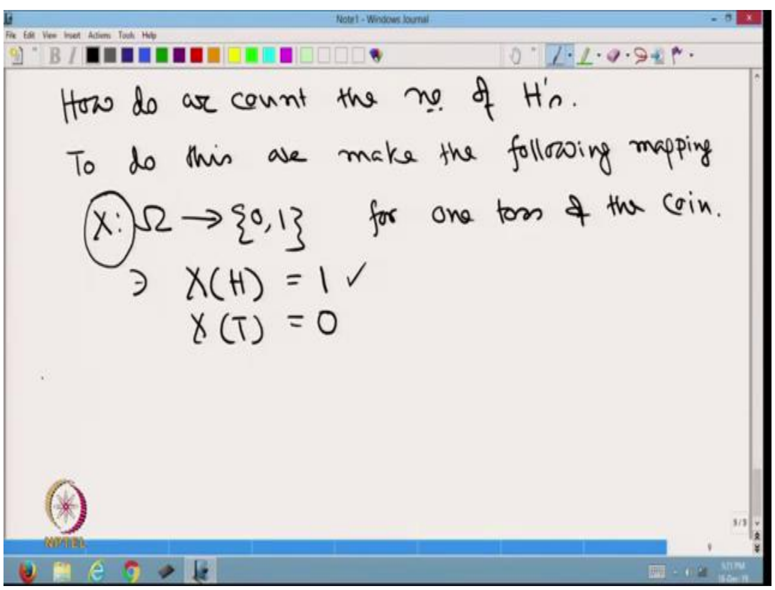
Similarly, if the item is defective, we may say we are getting a tail in the coin toss. Thus, when you look at 100 tosses, it is as good as checking hundred elements and if we count the number of heads, it is as good as counting the number of non-defective items in a batch of 100 items produced.

(Refer Slide Time: 12:48)



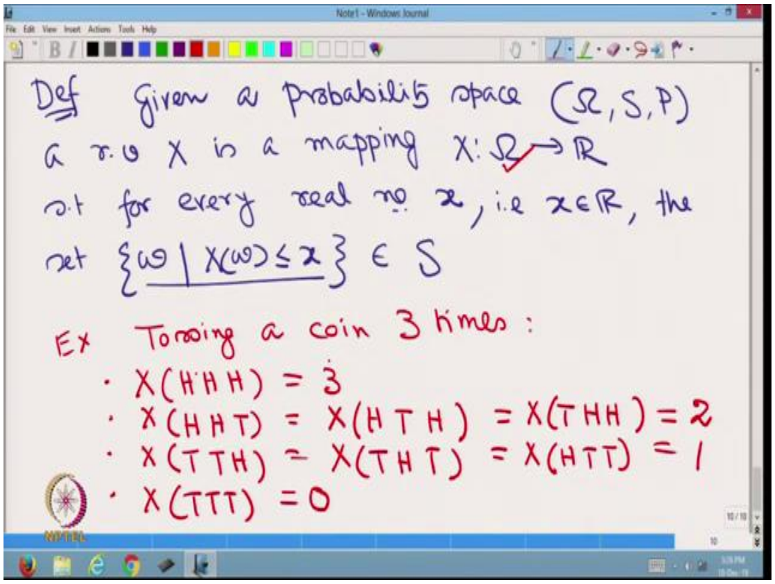
Therefore, it makes sense to map each sample point of the form say a HHTH… of length 100 as tossing a coin 100 times and counting of number of non-defective items or success is like counting the number of heads in that sequence.

(Refer Slide Time: 14:29)



Question is, how do we count the number of heads? Hence, we make a mapping X from omega to 0 1 for one toss of the coin. Such that, X of head is equal to 1 and X of tail is equal to 0. Therefore, corresponding to each outcome if we keep accumulating the values of H by collecting the number of ones at the end of the string, we shall know how many heads are there or how many non-defective items have been produced in this batch of 100 elements. So, this X, we can say is a random variable.

(Refer Slide Time: 16:17)

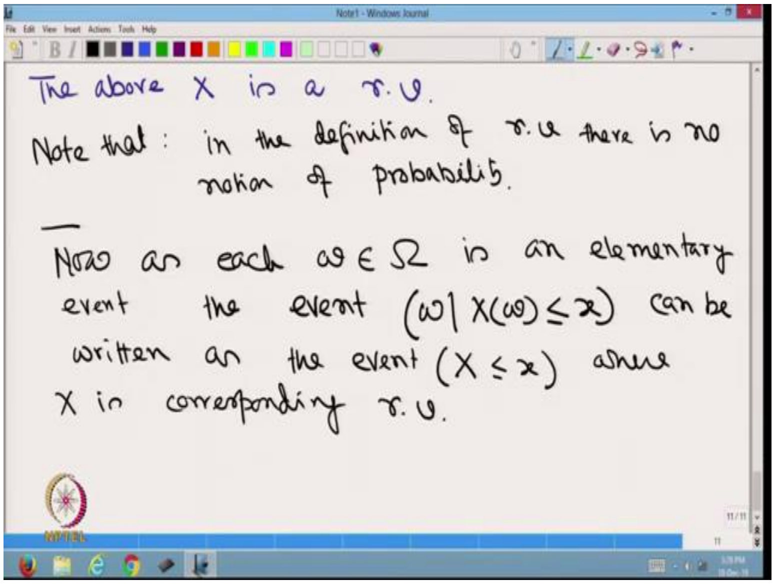


So, definition given a probability space omega, S, P. As you all know, omega is the set of possible outcomes, S is the sigma field defined over omega, and P is the probability measure. A random variable X is a mapping, X from omega to R such that for every real number x, that is x belonging to R the set omega such that X omega is less than equal to x belongs to S. That means, we look at the set of all sample points.

And for each one of them, we map it to the real number and such that for this set it has to be an element of that corresponding sigma filed. For example, tossing a coin 3 times, we have just seen the omega and we can write something like this X of HHH is equal to 3, because it is corresponding to three heads, X of HHT is equal to X of HTH is equal to X of THH is equal to 2.

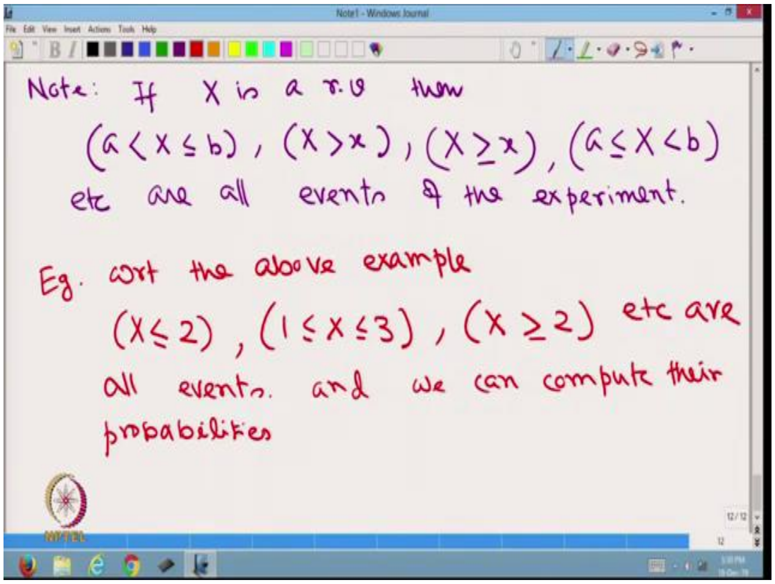
Because all of them are giving a two heads. X of TTH is equal to X of THT is equal to X of HTT is equal to 1, because, each of the sequence gives me one head and the X of TTT is equal to 0. Thus, we have mapped for each omega belonging to capital omega that is that set of possible outcomes, we are making a mapping from the W or the omega to the real numbers.

(Refer Slide Time: 20:18)



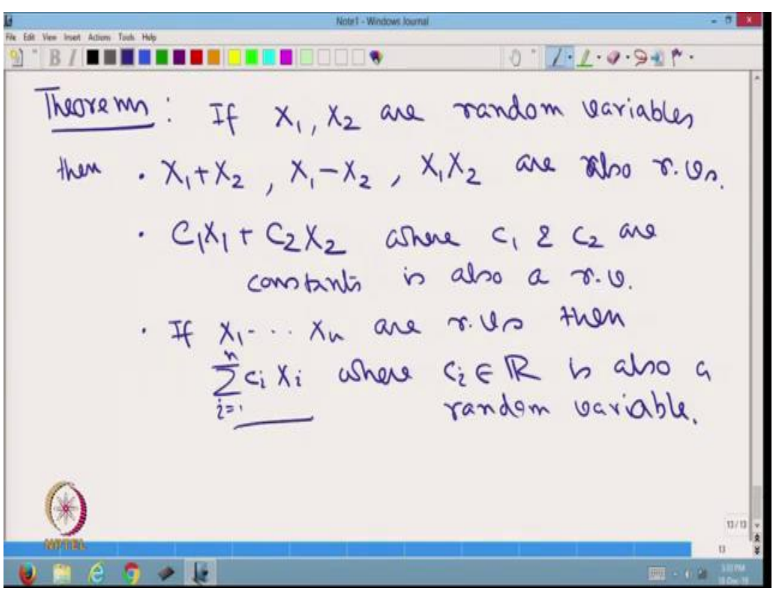
Therefore, the above X is a random variable. Note that, in the definition of random variable there is no notion of probability. So, when we talk about a random variable as a mapping, we are not discussing probability at all. Now, as each omega belonging to omega is an elementary event, the event omega such that X omega less than equal to x can be written as the event X less than equal to x where X is the corresponding random variable.

(Refer Slide Time: 22:13)



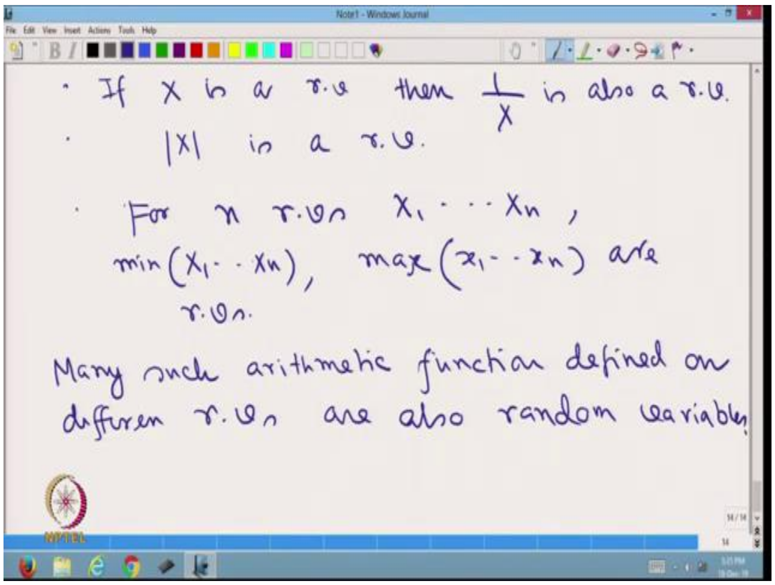
Note, if X is a random variable, then a less than x less than equal to b. X greater than x, X greater than equal to x, a less than equal to X less than b, etcetera are all events of the experiment. For example, with respect to the above example, X less than equal to 2, 1 less than equal to x less than equal to 3, X greater than equal to 2 etcetera are all events, and we can compute their probabilities.

(Refer Slide Time: 24:22)



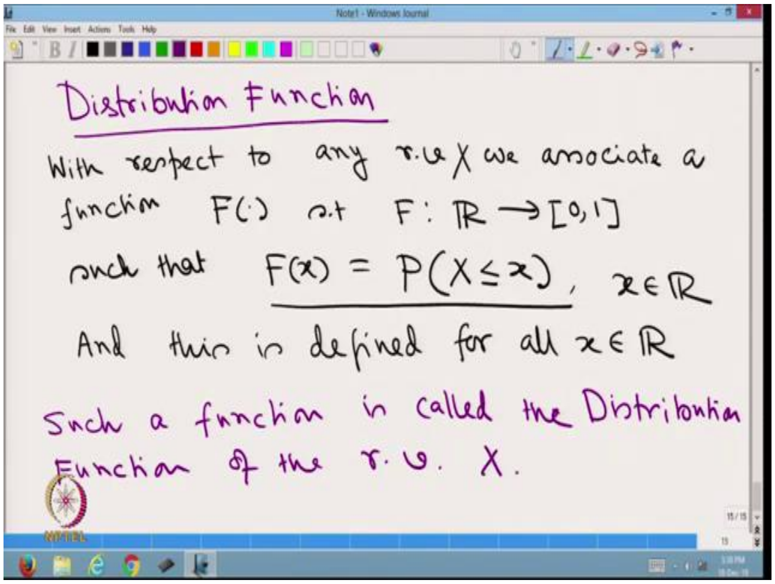
I now give you a few theorems, which I am not going to prove, but we should understand that if the X1, X2 are random variables then the next X1 plus X2, X1 minus X2, X1 X2 are also random variables. C1 X1 plus C2 X2 where C1 and C2 are constants is also a random variable. If, X1, X2… Xn are random variables then sigma Ci Xi, i is equal to 1 to n, where Ci belongs to R is also a random variable. That means, I am looking at a weighted sum of different random variables and that also is a random variable.

(Refer Slide Time: 26:07)



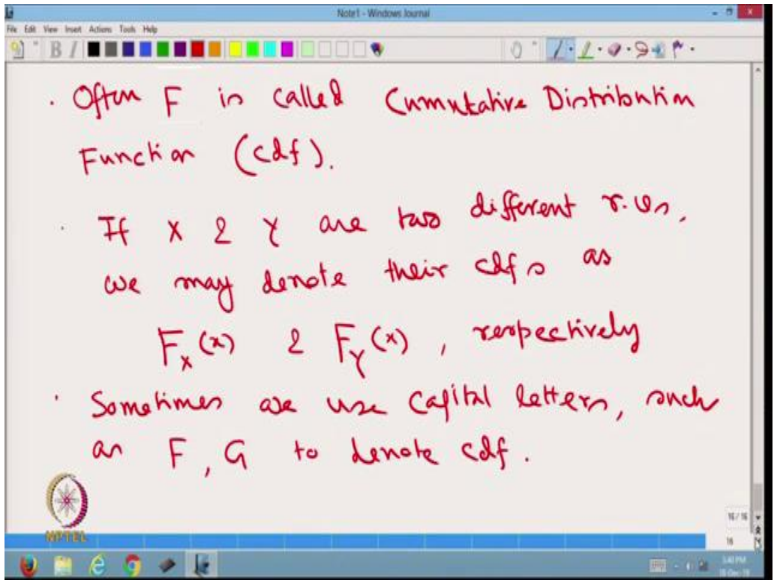
If X is a random variable, then 1 by X is also a random variable, mod of X is a random variable. For n random variables X1, X2… Xn, minimum of X1, X2… Xn, maximum of X1, X2… Xn are random variables, like that many such arithmetic functions defined on different random variables are also random variables. I am not going to prove this, but let us assume this results because proving this will need a lot of analysis, which is not a part of this course.

(Refer Slide Time: 27:49)



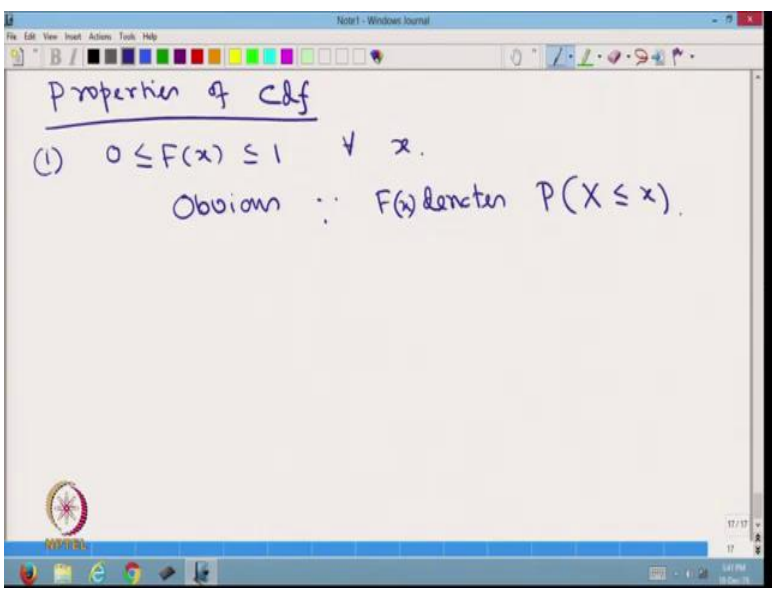
Distribution Function – With respect to any random variable we associate a function F such that F is from real line to 0, 1 such that F of x is equal to probability X less than equal to x where, so let me call it X. So, for each random variable x, we define the function F of small x, where X is a real number, such that F of x denotes that the probability that random variable X is less than equal to small x and this is defined for all X belonging to R. Such a function is called the Distribution Function of the random variable X.

(Refer Slide Time: 30:11)



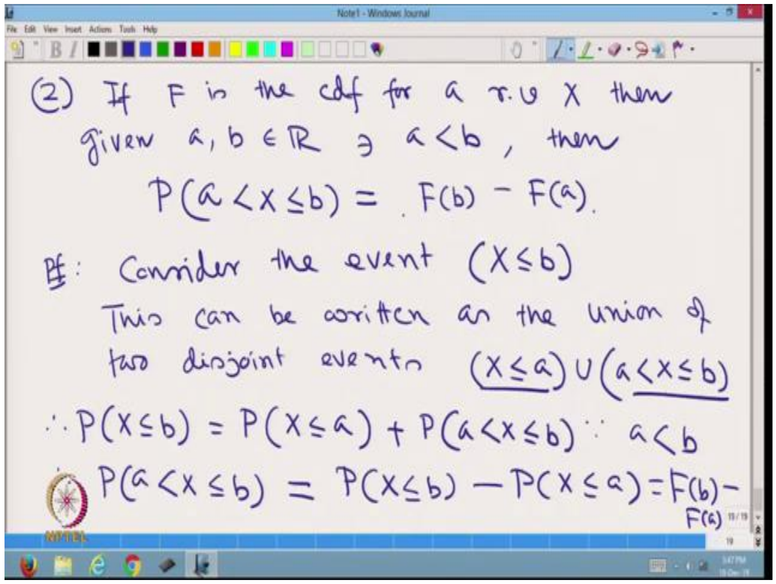
In some books, you may find F the term Cumulative Distribution Function or CDF to denote the same thing. If X and Y are two different random variables we may denote their CDF as FX x and FY x respectively or sometimes we use capital letters such as F, G to denote CDF, okay.

(Refer Slide Time: 31:53)



Properties of Cumulative Distribution Function. 0 less than equal to Fx less than equal to 1 for all x. This is obvious, since F denotes or Fx denotes probability x less than equal to x and we know that any probability will be between 0 to 1.

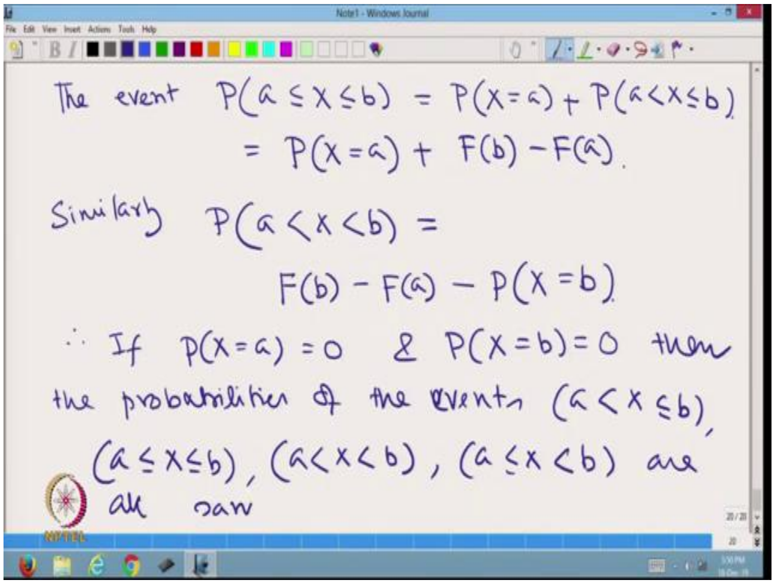
(Refer Slide Time: 32:46)



2 – If F is the CDF for a random variable X then given a comma b belonging to R such that a is less than b, then probability a less than x less than equal to b is equal to Fb minus Fof a. Proof: Consider the event X less than equal to b, that means, the random variable X taking any value, which is less than equal to b. This can be written as the union of 2 disjoint events X less than equal to a union a less than x less than equal to b, since a is less than b.

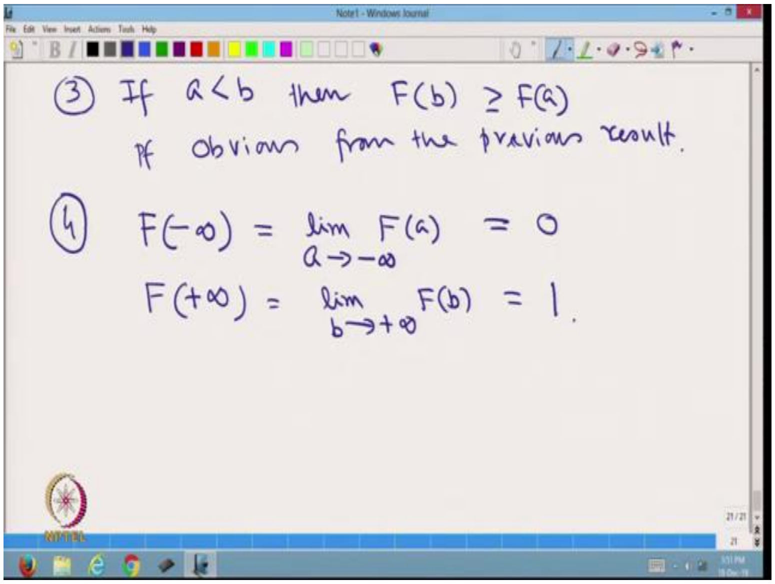
Now, since these two events are disjoint, therefore, probability X less than equal to b, can be written as the sum of these two individual events. Therefore, we can write it as probability X less than equal to a plus probability a less than X less than equal to b. Therefore, probability a less than X less than equal to b is equal to probability X less than equal to b minus probability X less than equal to a. And therefore, we can write it as Fb minus a Fa. So, that proves this result.

(Refer Slide Time: 35:41)



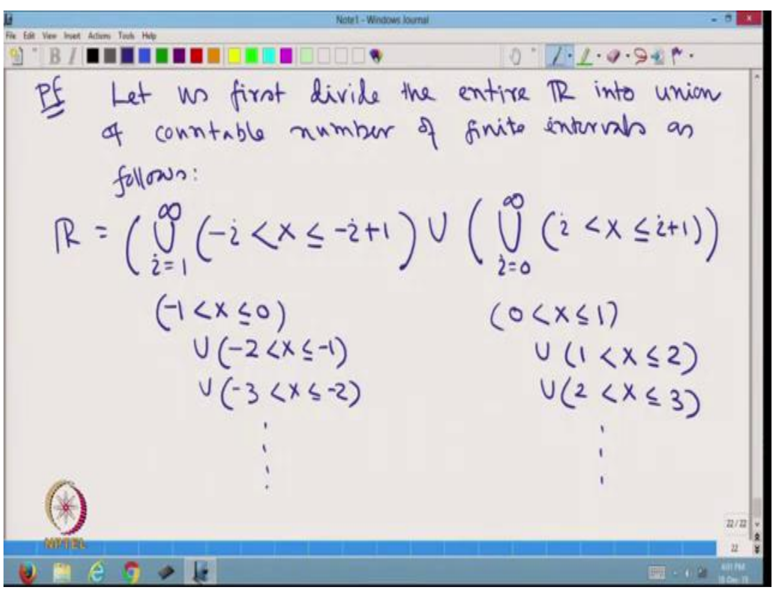
Therefore, the event, probability a less than equal to X less than equal to b is equal to probability X is equal to a plus probability a less than X less than equal to b is equal to probability X is equal to a plus Fb minus Fa. Similarly, probability a less than x less than b is equal to Fb minus Fa minus probability X is equal to b. Therefore, if probability X is equal to a is equal to 0 and probability X is equal to b is equal to 0, then the probabilities of the events, a less than x less than equal to b, a less than equal to x less than equal to b, a less than X less than b, a less than equal to X less than b are all same.

(Refer Slide Time: 37:44)



Next property, if a less than b, then F of b is greater than equal to F of a. Proof: Obvious from the previous result. 4 – F of minus infinity is equal to limit a going to minus infinity of F of a is equal to 0 and F plus infinity, which is equal to limit b going to plus infinity F of b, which is equal to 1.

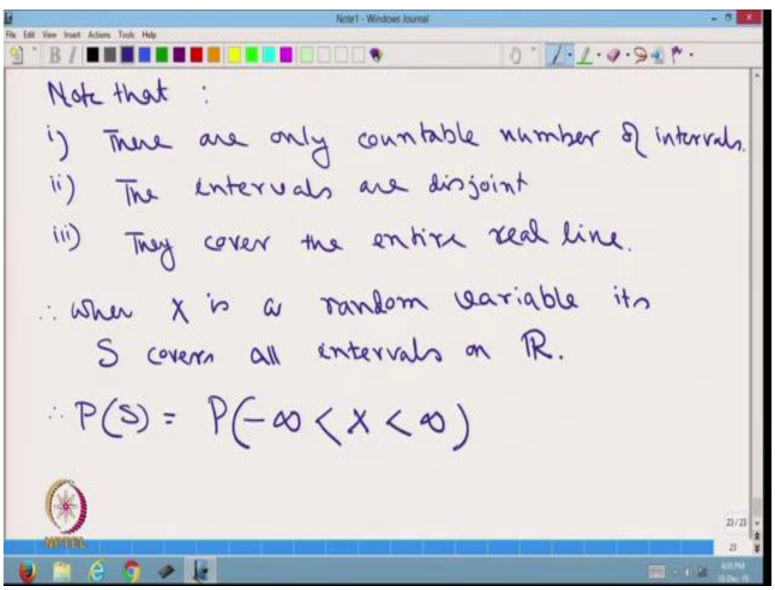
(Refer Slide Time: 38:48)



Proof: Let us first divide the entire R into union of countable number of finite intervals as follows: R is equal to Union over i is equal to 1 to infinity minus i less than X less than equal to minus i plus one union, union of i is equal to 0 to infinity, i less than X less than equal to i plus 1. So, basically what we are doing here we are looking at the set of intervals like this minus 1 less than X less than equal to 0 union minus 2 less than X less than equal to minus 1 union minus 3 less than X less than equal to minus 2.

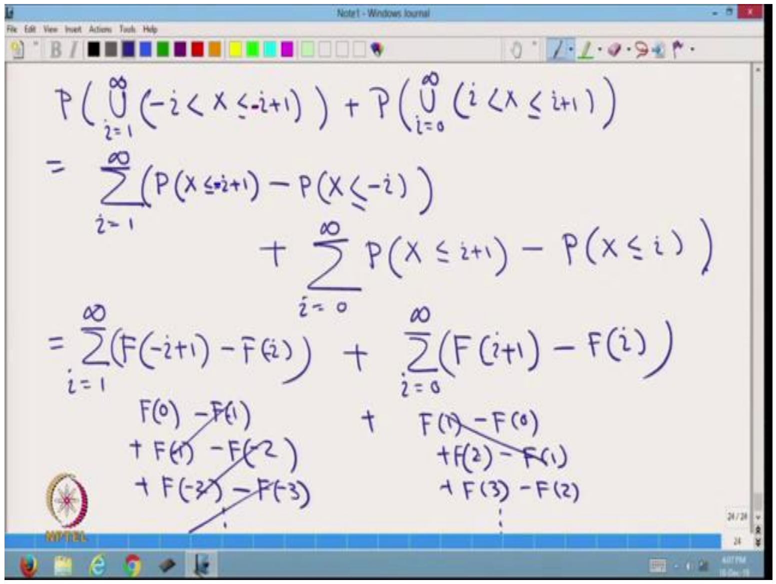
Like that, we are going to any minus n, n can be very large, but this does not matter, there will be an interval that is going to cover it. Similarly, on the positive side, we are looking at when i equal to 0 we get 0 less than X less than equal to 1 union when i is equal to 1, it is 1 less than X less than equal to 2 union 2 less then X less then equal to 3 like that, on the positive side, however large this number is, it is going to be a part of one of these intervals.

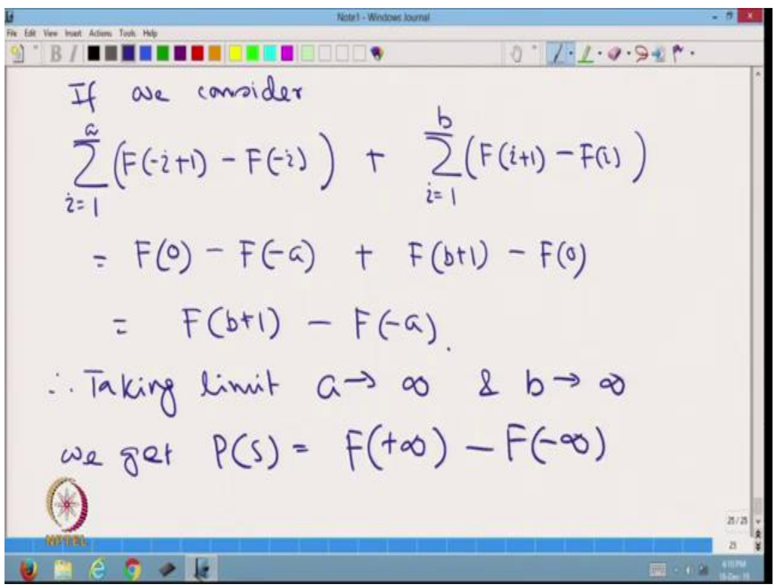
(Refer Slide Time: 41:30)



So, note that, 1 – There are only countable number of interval. 2 – The intervals are disjoint. And third, they cover the entire real line. Therefore, when X is a random variable its sigma field covers all intervals on R, therefore probability of the sigma fields is equal to probability minus infinity less than X less than infinity.

(Refer Slide Time: 43:14)



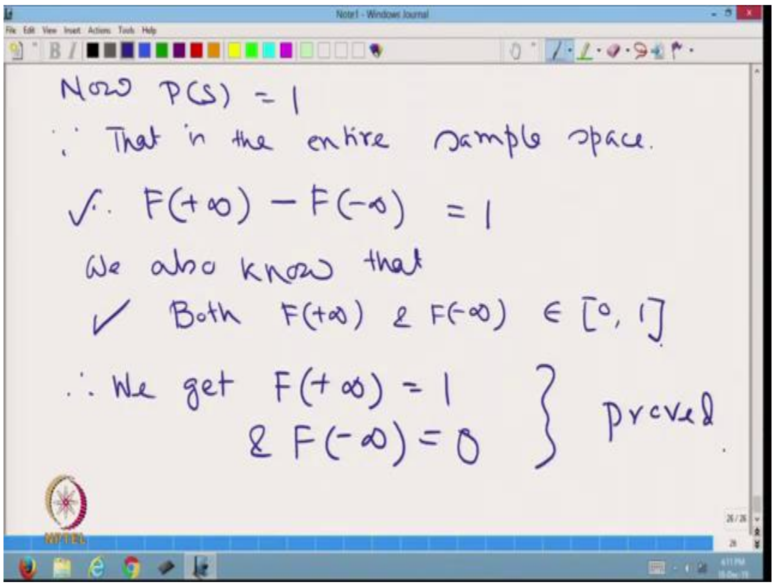


Which we can write now as probability union over i is equal to 1 to infinity minus i less than X less than equal to minus i plus 1 plus probability union over i is equal to 0 to infinity, i less than X less than equal to i plus 1 is equal to sigma i is equal to 1 to infinity, probability X less than equal to minus i plus 1 minus probability X less than equal to minus i plus sigma i is equal to 0 to infinity probability X less than equal to i plus 1 minus probability X less than equal to i.

Therefore, if we notice, this is going to be sigma i is equal to 1 to infinity, F of minus i plus 1 minus F of i plus sigma i is equal to 0 to infinity F of i plus 1 minus F of i. So, here we are getting terms like F of 0 minus F of 1 plus F of 1 minus, this is minus, so this is minus F of minus 1 minus F of minus 2 plus F of minus 2 minus F of minus 3, like that. And here what we are getting F of 1 minus F of 0 plus F of 2 minus F of 1 plus F of 3 minus F of 2 like that, we are getting a sequence of terms. Note that, they keep on canceling each other. Similarly on this side, they keep on canceling each other.

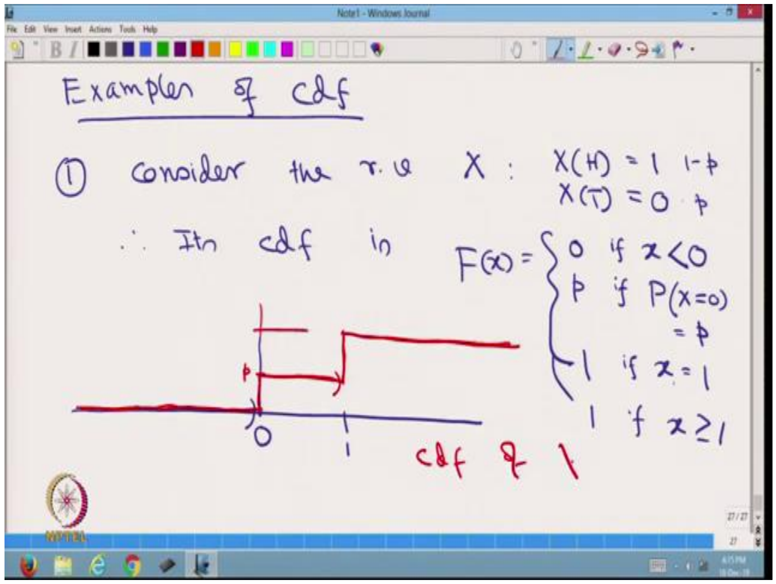
Therefore, if we consider sigma i is equal to 1 to a F of minus i plus 1 minus F of minus i plus sigma i is equal to, 1 to say b, F of i plus 1 minus F of i, then after cancellation, what we are going to have is equal to everything will cancel. So from here, instead of everything we are going to get is F0 minus F of minus a and from here we are going to get F of b plus 1 minus F0 is equal to F of b plus 1 minus F of minus a. Therefore, taking limit a goes to infinity and b goes to infinity, we get P of S is equal to F of plus infinity minus F of minus infinity.

(Refer Slide Time: 48:18)



Now, P of S is equal to 1, since that is the entire sample space. Therefore, F of plus infinity minus F of minus infinity is equal to 1. We also know that both F of plus infinity and F of minus infinity belong to 0, 1. So, to satisfy both of these, we get the F of plus infinity is equal to 1, and F of minus infinity is equal to 0, proved.

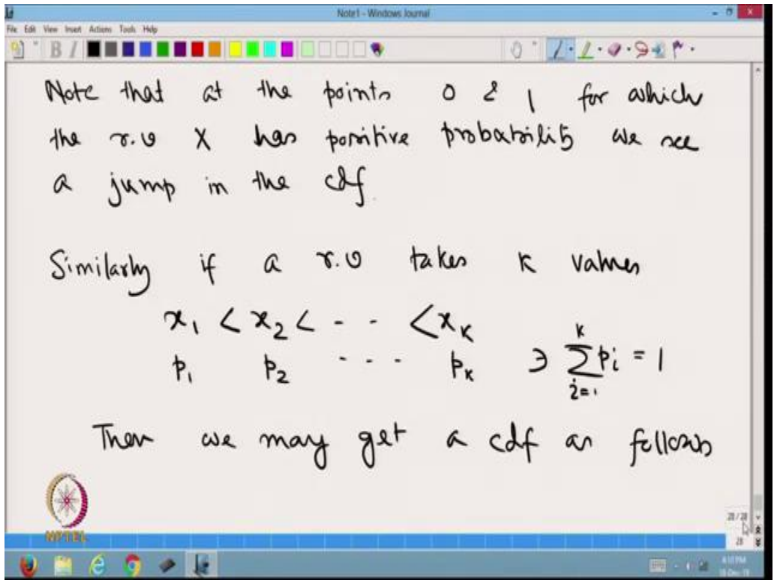
(Refer Slide Time: 49:33)



So, let me give you some examples. Consider the random variable X such that X of H is equal to 1 and the X of T is equal to 0. Therefore, its CDF is F of X is equal to 0 if X less than 0, this is equal to p if probability X is equal to 0 is equal to p, this is equal to 1, if x is equal to 1 and it is 1, if x greater than equal to 1. So, let me explain that suppose this is the entire real line, this is 0, this is 1 random variable takes only two values 0 and 1, we did not assign any probability. So, suppose probability X is equal to 0, that probability is p and XH is equal to 1, therefore, this is going to be 1 minus p.

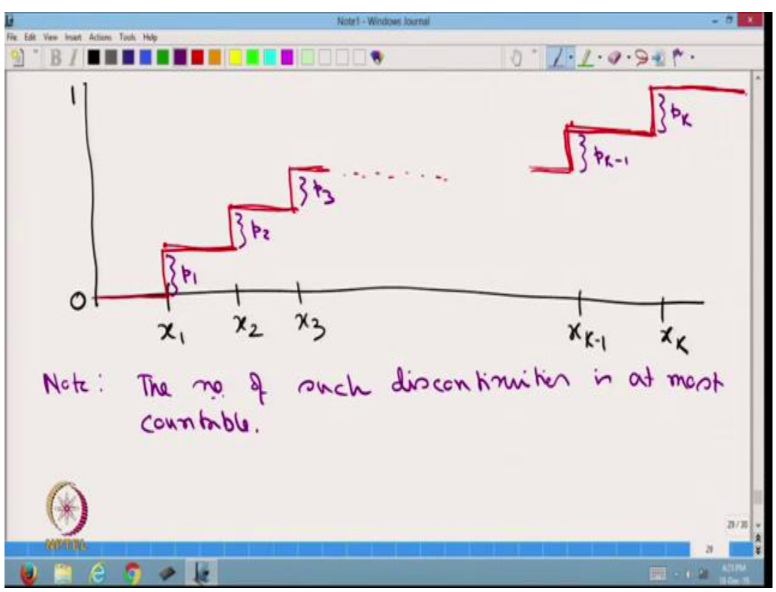
Therefore, for any X, this value is going to be 0. When X is equal to 0, the probability of F 0 is equal to p therefore, at the point 0, if we consider, what is the probability that X less than equal to 0? That value is going to be p and that will continue for all values less than 1. Now, at 1, the probabilities 1 minus p therefore, X less than equal to 1, that probability is going to be 1. Therefore, at this point, it is going to jump to 1 and after that it will continue like this. So, this is the graph of the CDF of X.

(Refer Slide Time: 52:47)



Let me now give you another example. Note that, at the point 0 and 1 for which the random variable X has positive probability we see a jump in the CDF. Similarly, if a random variable takes k values x1 less than x2… xk such that the probability of x1 is P1, probability of x2 is P2, probability of xk is equal to Pk such that sigma pi, i is equal to 1, 2 k is equal to 1, then we may get a CDF as follows.

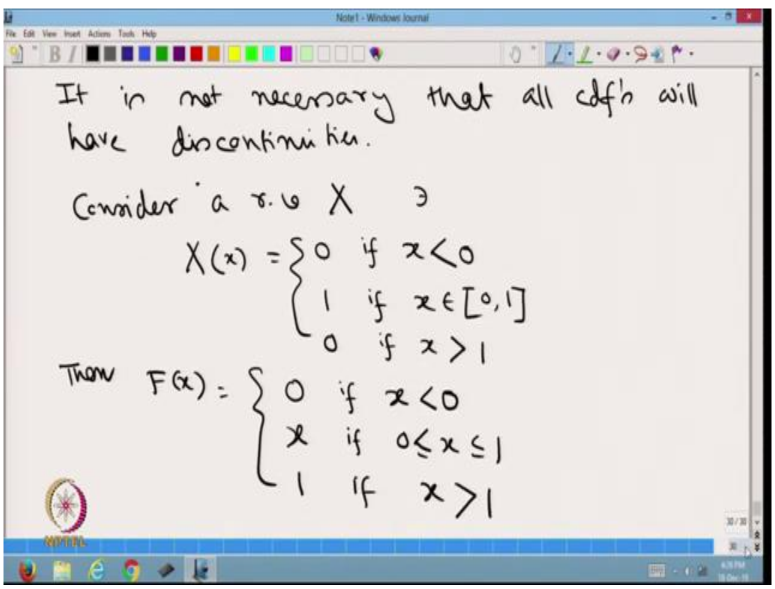
(Refer Slide Time: 54:32)



So, if this is the real line and these are the discrete points x1, x2, x3… xk minus 1, xk suppose, this is 0 and this is 1, then we would expect a CDF which will look like this at x1, there will be a jump, at x2 there will be the next jump like that at x3 there will be another jump, at xk minus 1 there is going to be a jump and finally, at xk there is a jump and the value takes 1.

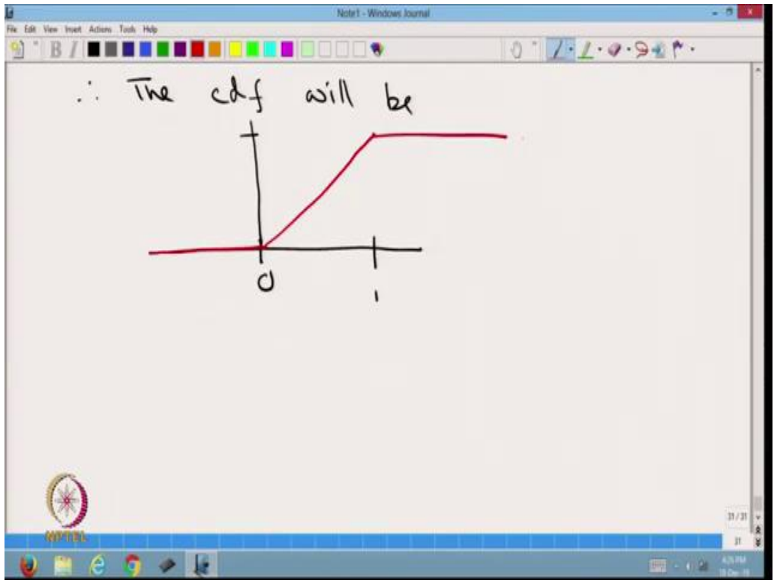
Thus, we get something like a step function which looks like this, I hope you understand why such a shape will come. Now, note that the number of such discontinuities is at most countable, that is, there cannot be unaccountably many discontinuities in the CDF of any random variable. And I hope you could guess the quantum of jumps. This is the P1, this is P2, this is P3, this is Pk minus 1, this is Pk. And finally, it is going to be 1 because the sum of all the probabilities is equal to 1.

(Refer Slide Time: 57:01)



It is not necessary that all CDFs will have jumps or discontinuities. For example, consider a random variable X such that X of x is equal to 0 if x less than 0 is equal to 1 if x belongs to 0, 1 and it is 0 if x is greater than 1. Then F of x is equal to 0 if x less than 0, this is equal to x, if 0 less than equal to x less than equal to 1 is equal to 1, if x greater than 1.

(Refer Slide Time: 58:34)



Therefore, the CDF will be, thus, there is no discontinuity and F of minus infinity is equal to 0 F of infinity is going to be 1 and between 0 to 1, it is monotonically increasing in a linear fashion. Try to convince yourself that you are going to get this CDF for the corresponding random variable. Okay friends, I stop here today. In the next class, I shall introduce different discrete random variables, which are very important from not only different applications. Thank you.