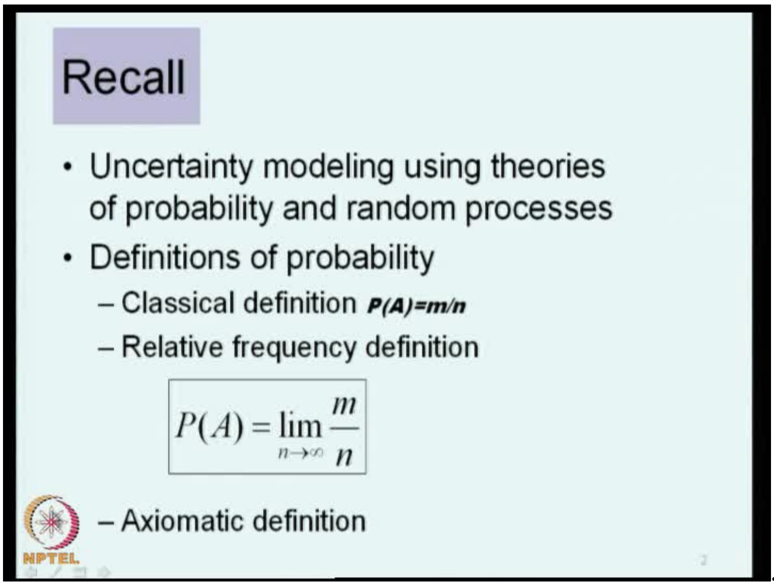
Stochastic Structural Dynamics
Prof. Dr. C. S. Manohar
Department of Civil Engineering
Indian Institute of Science, Bangalore

Lecture No. # 02
Scalar Random Variables-1

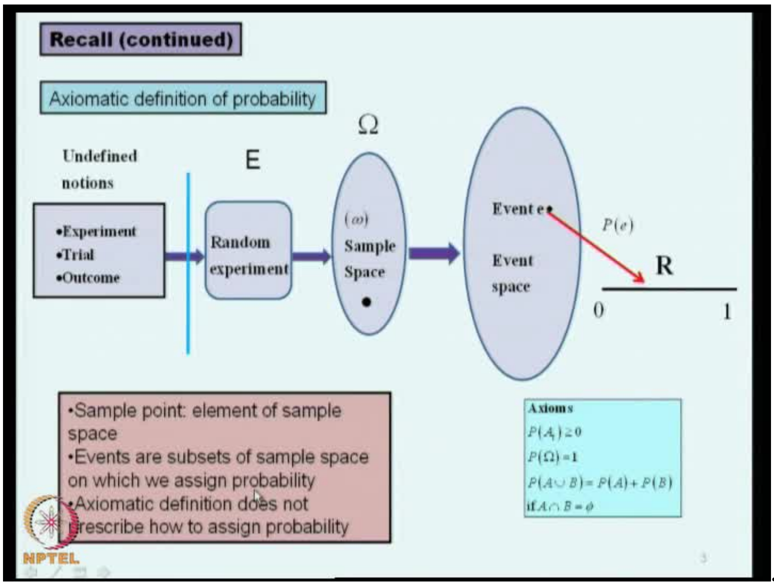
(Refer Slide Time: 00:22)



Before we begin the second lecture, we will quickly recall what is that we did in the previous lecture. We agree to model uncertainties in structural engineering problems, using theories of probability and random processes, and with that in mind, we started defining what the word probability means. There were three alternative definitions that we considered; one is the classical definition, where probability of an event A is defined as ratio m by n, where n is the number of possible outcomes of a random experiment and m are the number of events favorable to the event A. This assumes that all outcomes of a random experiment are equally likely and there was no room for experimentation in this.

The next definition, the relative frequency definition, based purely on experimentation; here, if a random experiment is performed n number of times, and out of these n trials, if m of them are favorable to event A, then probability of A is defined as the ratio m by n, where these number of trials is taken to be sufficiently large. The problem with this definition, is that, it does not allow us to assign probability is to situations where we cannot do an experiment, and when the meaning of the statement, that n is sufficiently large is not mathematically very precise. So, this led to the definition of more acceptable definition known as axiomatic definition.

(Refer Slide Time: 02:05)

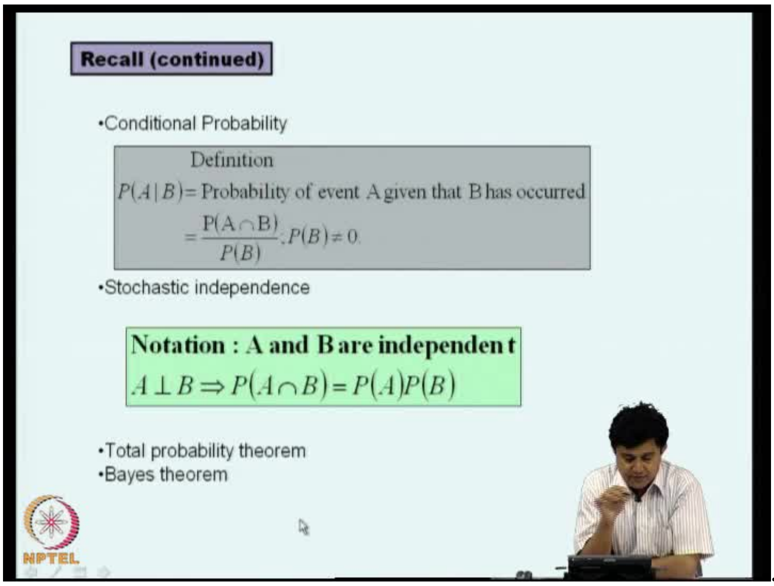


Here, what we did was, we started with certain notations which we considered as primitives, in the sense we do not define them, terms like experiment, trial, outcome etcetera, we do not offer a definition. So, experiment is the physical phenomena that could be observed repeatedly, and a single performance of an experiment is a trial, and the observation we make on performing a trial, we call it is an outcome.

So, with this description, we introduce the notion of a random experiment, which is an experiment in which we know, what all the possible outcomes of a trial are, but in any given trial, we do not know what exactly will be the outcome. Now, the set of all possible outcomes of a random experiment, we call it as sample space; this sample space itself could be finite, countable infinite or uncountable infinite. We would like to studies subsets of this sample space. And therefore, we define another larger set known as event space; this is collection of subsets of omega, this sample space, on which we would like to assign a measure known as probability which obey this axiom, that is, axiom of positivity axiom of normalization and axiom of additivity.

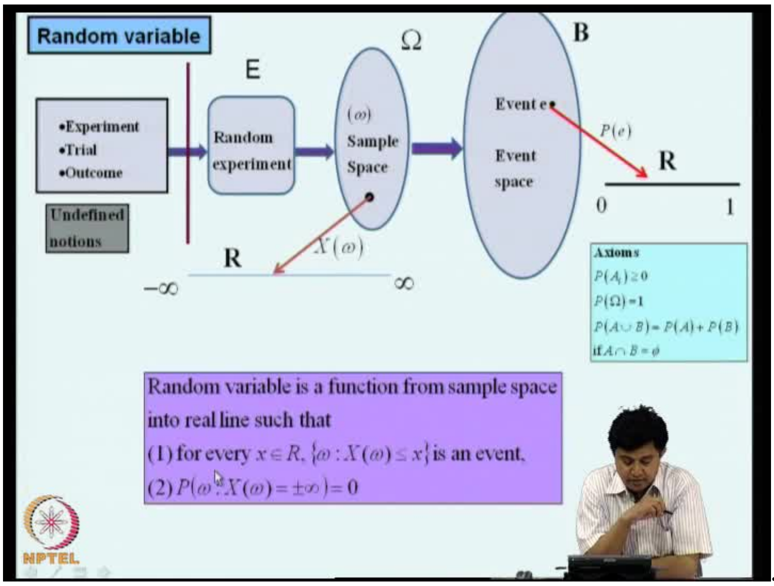
Now, a point in sample space we call it as sample point, and the term events are taken to means subsets of events space on which we assign probability obeying the rules of, I mean, these axioms. Now, one important thing we should notice here, is that, the axiomatic definition of probability does not prescribe how the probability should be measured; what it is simply says is, if you are going to assign a probability to an event, obey these rules. It is like, analogy would be like in geometry, we simply state theorems you should be obeyed by lines and curves, and so on, so forth, we never talk about how to measure length of a line or a curvature of a surface and the actual measurement issues are not addressed.

(Refer Slide Time: 04:35)



The question of how to prescribe probability belongs to the subject of statistics, and at a later point in the course, we will return to this question. Following the definition of probability, we introduce the notion of conditional probability, where we define probability of an event A conditioned on B, that means, probability of event A given that B has occurred, is by definition, this ratio, probability of A intersection B divided by probability of B and probability of B should not be equal to 0. The notion of conditional probability, based on this notion will defined, what is meant by independence of two events? We say that, events A and B are independent, if probability of A intersection B is probability of A in to probability of B.

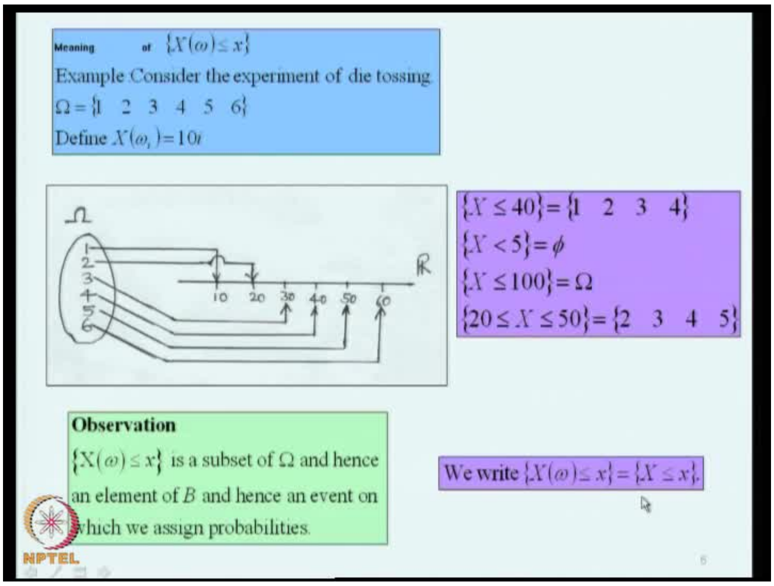
(Refer Slide Time: 05:36)



This would also mean that, probability of A is not affected by the occurrence of event B. We briefly talked about total probability theorem and Bayes theorem, there will be occasions for us to use this; so, I would not like to elaborate on this at this stage. In today's lecture, I would like to introduce the notion of a random variable that is what we will begin today. Again we return to this scheme, schematic description of axiomatic definition of probability, this is an event space, this is sample space; this is the random experiment, sample space, event space, and the probability.

Now, we define a function, random variable is a function, that is define, from sample space into the real line, such that, that means, for every element in sample space, we assign a number on the real line, such that the probability, we would like to consider the set of all omega; omega is the point in sample space such that X of omega less than or equal to X is an event, in the sense, for this set we would like to assign a probability; so, I need to clarify what this set means. We also impose certain restrictions stated here.

(Refer Slide Time: 06:57)



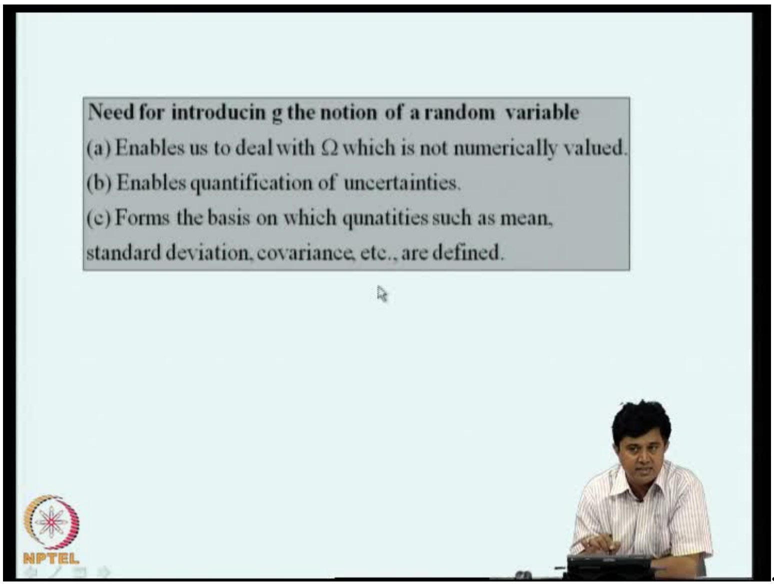
Now, let may clarify what is the meaning of this set; so, to do that, we will consider a simple example of die tossing. So, the sample space here is 1, 2, 3, 4, 5, and 6, this could as well being red, yellow, blue, or any 6 colors as well, but in, so happens that I am on the die, we are seeing those dots and I am counting the number of dots. I define now, a function X of omega i as 10 I, where i is 1, 2, 3, 4, 5, 6; that means, if this is a sample space, I map 1 to the point 10, 2 to the point 20, and so on, and so forth 6 to the point 60.

So, this mapping defines the definition of random variable, and this provides a definition of random variable. Now, let us now look at, what is the meaning of the set X of omega less than or equal to X? We are temptate to interpret this as a part of the real line, but that is not what exactly we mean; we would like to interpret this as a subset of omega, and therefore, an element in B, because this is a finite sample space and which we would like to assign probabilities. Let us consider, what is X less than or equal to 40? So, 40 is here. What are the events, what are the points in sample space which lie to the left of these points? That is 1, 2, 3, 4; the subset of sample space namely 1, 2, 3, 4, is this set, it is not the interval from minus infinity to 40, you should not interpret that way; it is actually subset of sample space where elements in that or map to the left of this point.

Similarly, what is X less than or equal to 5? So, you look at 5, at, or there any points in sample spaces which are map to this line? No, therefore, this is a null set. Now, you look at X less than or equal to 100, so 100 is somewhere here, here, and we see that all these points 1, 2, 3, 4, 5, 6, are map to the left of that point; so, therefore, X less than or equal to 100 is nothing but the sample space itself.

Now, one more example, 10 will not be in that, 20 will be in, 30 will be in, 40 will be in, and 50 will be in; so, that means, 2, 3, 4, 5. Now, this set X of omega less than or equal to X, for sake of simplicity, we write it as X less than or equal to X, we omit this omega; omega is a sample point, point in the sample space that we will not write explicitly as a convention.

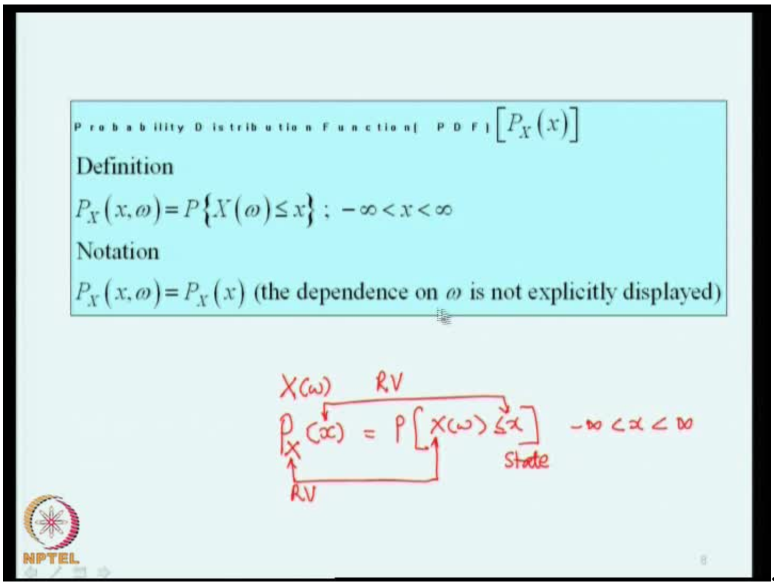
(Refer Slide Time: 10:16)



Now, what is the need for introducing the notion of a random variable? It will become clear as we go along, but at this stage, first order description of that would be, that it enables as to deal the sample space which is not numerically valued; for example, you toss a coin, you get head or tail, they are not numbers; so, you can as well say, when I get head, I will call it as 0, and when I get tail I will call it as 1, that is the definition of random variable for that random experiment. By talking about numbers, we immediately would be able to quantify the uncertainties; instead of talking about head and tails, since I am talking about numbers, I would be able to quantify uncertainties in some sense.

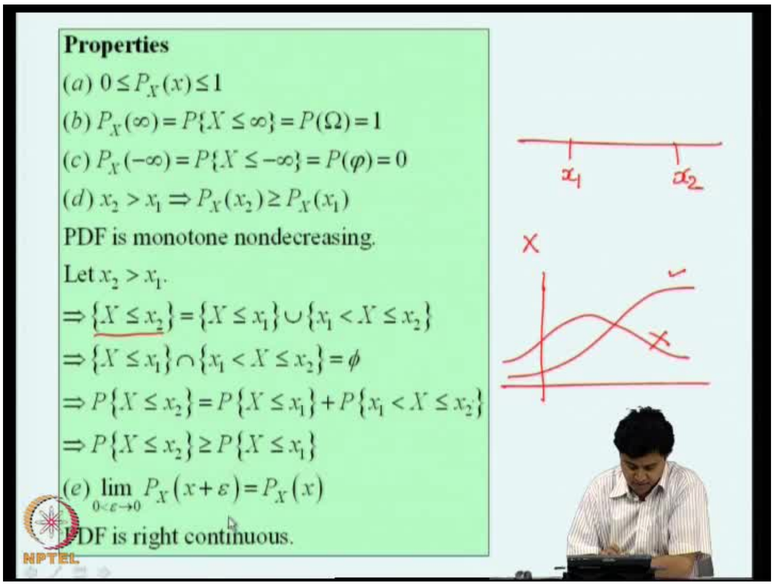
You would also see later that the introduction of random variable, also forms the basis on which we talk about descriptions, such as mean, standard deviation, covariance, and so on and so forth, which we made a heard about, but will now make the meaning of these term precise in due course.

(Refer Slide Time: 11:36)



Now, we introduce a function known as probability distribution function, the notation we use is P x of x. Now, X of omega is the random variable, the probability distribution function is notation is this, that is what I am following, this as we writing here, this is probability of X of omega less than or equal to X, where x itself takes values from infinity this X, that I am writing as the subscript refers to this random variable; this lower case x, I am writing refers to this state; this is the random variable. This function as we vary x, if we were to plot P x of x versus x that function is known as probability distribution function; again, this omega we were not going to write when we write probability distribution function, that is, the dependence on omega is not explicitly displayed.

(Refer Slide Time: 13:16)

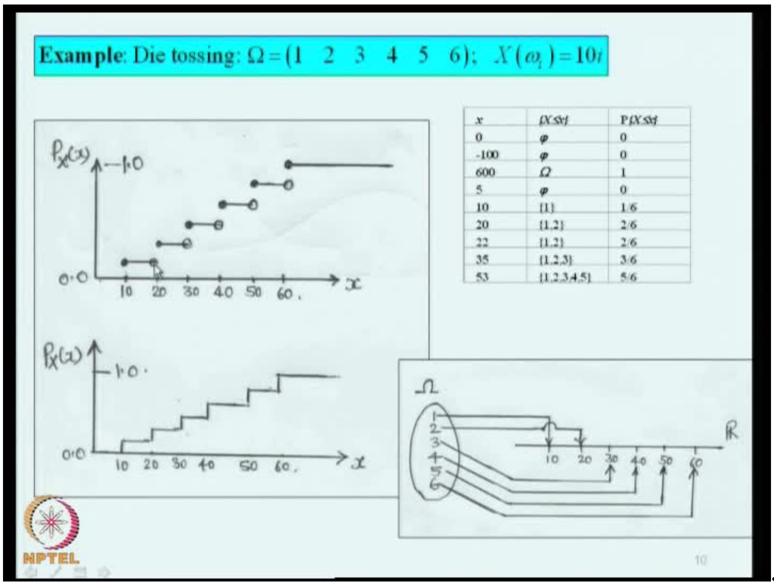


Now, what are the properties of this probability distribution function? It is essentially probability, therefore it should be a number between 0 and 1, it cannot be negative, it cannot be greater than 1. P x of infinity is what? X less than or equal to infinity; X less than or equal to infinity is a nothing but the sample space, probability of sample space is 1 because that is the axiom. P x of minus infinity is what? Is the probability of X less than or equal to minus infinity, the, this event X less than or equal to minus infinity is nothing but a null set; probability of null set is 0 which is again one of the corollaries that we saw in the previous lecture.

An important property of probability distribution function, is that, it is monotone non-decreasing, that means, as X is increased, the probability distribution function can remain constant or it can increase, it cannot reduce why is that so? If we consider two points x 1, x 2, the event X less than or equal to x 2, can be written as X less than or equal to x 1 union x 1 X greater than x 1 and less than or equal to x 2; these two events are mutually exclusive. Therefore, the probability of X less than or equal to x 2, can be written as probability of X less than or equal to x 1 plus the probability of this event.

This probability itself is a number between 0 and 1, consequently if I drop this term here, they, it would mean probability of X less than or equal to x 2, at the most can be equal to probability of X less than or equal to x 1 by monotone non-decreasing, this is what I mean. So, probability distribution, typical probability distribution function can be like this, this is, but it cannot be like this.

(Refer Slide Time: 15:50)



Now, there is one more property that we need to understand, I will elaborate this shortly, the probability of density function is written continuous; I will explain that slightly later we will continue now. By considering an example, we again written to this example of die tossing; this coin tossing and die tossing experiments are time onward simple experiments to learn about probability that is why I am using them. So, we again consider this die tossing experiment, the sample space is 1, 2, 3, 4, 5, 6 the random variable, we defined as X of omega as 10 i.

So, now, let us plot the lower case x here, on the x axis here and go and assigning values to x; suppose X equal to 0, what is X less than or equal to 0? It is a null set; therefore, probability of X less than or equal to X is 0, that means, at 0 I should plot 0; however minus 10 X less than or equal to minus 100 is phi, null set probability continues to be 0.

Now, what happens at 600, probability X less than or equal to 600 is probability of the sample space, therefore that is 1; that means, this, here it is 0, and as X becomes large, it is becoming 1. Now, how about 5? It is still 0, how about 10? X less than or equal to 10 is we capture the first event, here getting 1, and if the die is fair the that probability will be 1 by 6 will use the classical definition; so that, at X equal to 10, therefore the curve jumps from 0 to 1 by 6. how about it 11, X equal to 11, so 11 is here, so the only event is

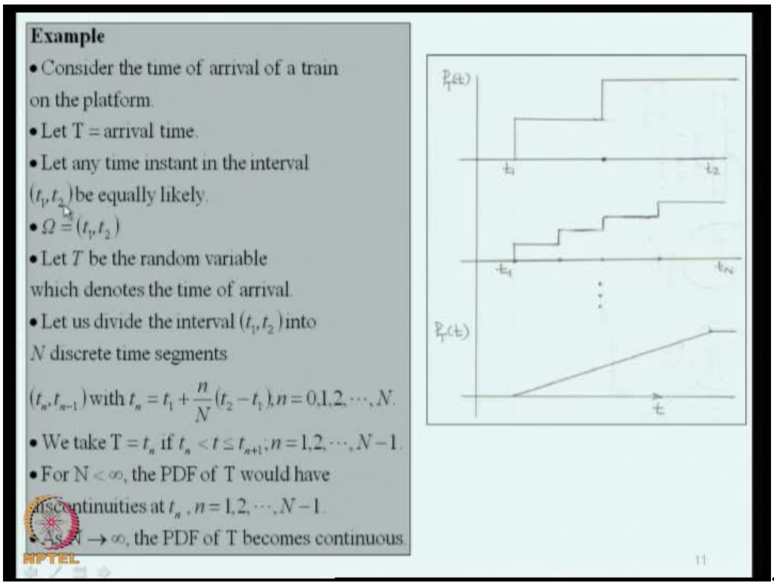
10, only event is 1, the probability therefore continues to remain constant till we encounter the point 20; at 20 what happens? X less than or equal to 20, so in that bus gate, there will be one more event 1, 2.

So, these getting 1 and getting 2 are mutually exclusive; therefore, union of 1 , 2 is probability of 1 plus probability of 2; so, I will get 2 by 6, let me from 1 by 6, it now jumps to 2 by 6. So, as we continue you can assign 20, 22, 35, let us say for example, 35, we are here; so, there are three sample points 1, 2, 3, therefore probability is 3 by 6, so at 35 again 3 by 6. Now, you notice now I am putting a dark spot here and a circle here, at x equal to 20, the probability distribution function is discontinuous. As a convention, whenever there is a discontinuity, we would report the value of probability distribution function from the limit from the right hand side, that means, probability distribution function at 20, we should report it as 2 by 6, there is the convention again.

So, this is actually the probability distribution function, you can easily see, now all these numbers are between 0 and 1 and it is monotone non-decreasing, there is no, nowhere the probability distribution function actually comes down. Now, in further representations we will not bother to write this dot and these circles, eventually will stop writing that, if we are to do that you will plot the same graph in this manner, it is like a staircase function; this is the probability distribution function of this random variable.

Now, if we are given this picture, we should notice that whatever we need to learn about this random experiment is encapsulated in this graph, that will become again clear as we go along you should notice that right away.

(Refer Slide Time: 19:51)



Now, one more experiment, let us we consider the experiment of arrival of train on a platform; so, let capital T be the arrival time, let any time instant in the interval t 1 to t 2 be equally likely l; so, the random experiment here is (( )), when the train arrives on the platform. Now, whatever, therefore, the sample space is the interval t 1 to t 2, so it will not arrive before t 1 and it would have certainly arrived after t 2.

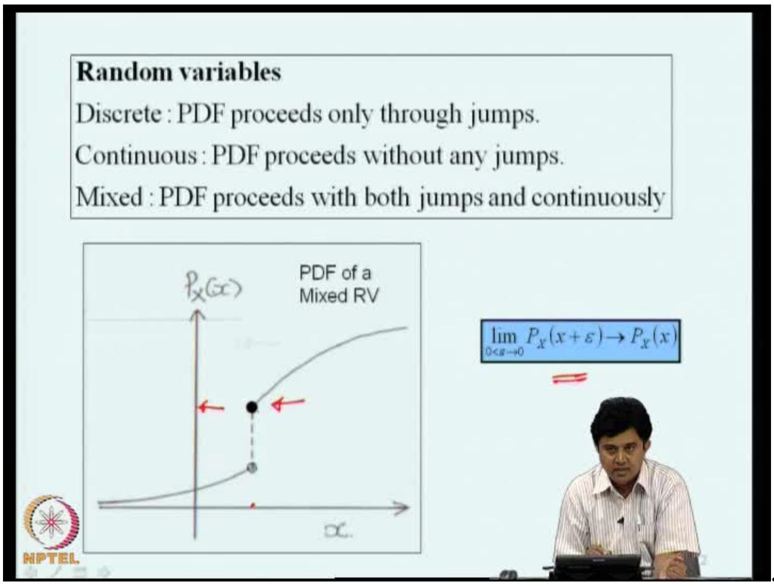
Now, let T be the random variable which denotes the time of arrival, so t belongs to the interval t 1 to t 2. Now, we will begin with the simple model where the interval t 1 to t 2 is divided here; t 1 to t 2 is divided into two intervals, for example, this is 10 o'clock, this is 11 o'clock, I will say 10 to 10.30 is my slot 1, 10.30 to 11 is my slot 2, so what is the probability that train would arrive between 10 to 10.30? Is 0.5, so this is the 0.5 R I t.

Now, what is the probability that capital T is less than this point, it is train arrival in slot 1 and train arrival in slot 2; therefore, the probability is 1. If you are to make 4 slots, you will get number of… We will now consider another example, the random experiment here is as follows, consider time of arrival of a train on a platform, so let us denote by capital T, the arrival time; let us assume that the train arrives on the platform, any time, during the interval t 1 to t 2; so, the sample space here is the interval t 1 to t 2, the random variable is the time at which the train arrives that itself is the number; so, it is t 1 to t 2.

Now, so, T be the random variable which denotes the time of arrival. Now, what will do is, we will now focus on this sketch, I am dividing the interval t 1 to t 2 into two intervals, two sub intervals, and now I will consider the random experiment, where I will observe, if train has come in slot 1 or slot 2, for example, this is 10 o'clock and this is 11 o'clock, the first slot is 10 to 10.30, second slot is 10.30 to 11, now what is the probability that train arrives would have arrived at say 10.15, is 1, is half, because the probability that the train would arrived between 10 and 10.30 is half; so, this is half. So, if you consider a time instant between 10.30 and 10.45, the probability will be 1.

So, here the sample space has two outcomes; now, the interval t 1 to t 2, can now be divided in to 4 such slots, here this jump in probability distribution function is 1 by 4. So, as we divide this interval into larger number of sub intervals in the limit, for example, if I take n discrete time segments, then we can take capital T as t n, if the arrival is between t t n to t n plus 1; so, typically the probability distribution function of this discrete random variable would be a stair case of n steps, at each step it will increment by t 2 minus t 1 by n; now, in the limit of this n becoming infinity, you would see that this staircase would now become a ramp and this probability distribution function now corresponds to a random variable which we call as being continuous.

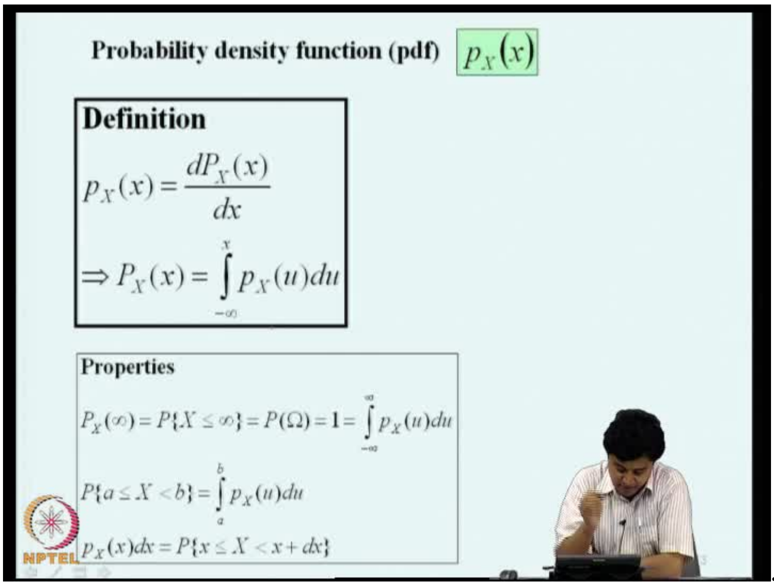
(Refer Slide Time: 24:27)



This is the continuous random variable; the sample space here is the real line t 1 to t 2. Thus, we can now see that we can classify random variables in to three categories: one is discrete random variables, other one is continuous random variables and the third category known as mixed random variables. A working definition for these random variables could be this, in a discrete random variable, the probability distribution function proceeds only through jumps, any increasing probability distribution function is only through a jump; where as in a continuous random variable, the probability distribution function proceeds without any jumps; where as in a mixed random variable the probability distribution function proceeds with both jumps and continuous segments, for example, this is the probability distribution function of a mixed random variable; so, in this segment, it is increasing continuously, and at this point, there is a discrete jump, and then it continuous to increase continuously.

So, in a general type of random variable, there could be several such jumps and there could be several such continuous segments. Now, again let us written to the point where there is a jump; you see here, I have drawn a circle here and a blob here; so, if I want to know the probability of distribution of the value of the probability distribution function, at this point, this is the definition that we will adapt. We will take a point which is x plus epsilon where epsilon is positive, that means, we approach from this side. And this is the number, that we report as probability distribution function at this value; all though, this also a, I mean, there is a discontinuity here, if you approach from left you will get this, right you will get this, but as a convention we assume that probability distribution function is write continuous.

(Refer Slide Time: 26:34)

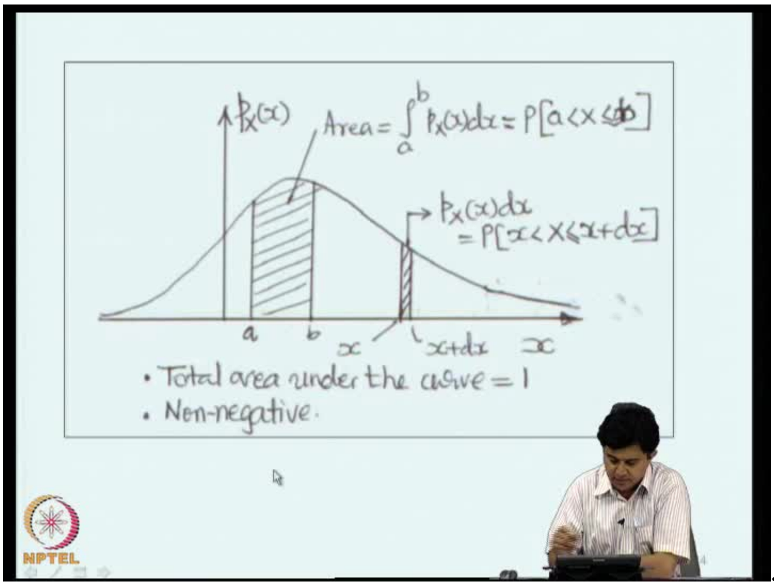

.

We now introduce another description of a random variable namely, probability density function; this is written with a lower case p, this subscript is again a capital letter X and the argument is a lower case x. The definition is the probability density function of random variable X evaluated at, the probability density function of random variable X evaluated at this lower case x, is actually the slope of the probability distribution function at x; it is by definition the derivative of the probability distribution function.

So, from this you get probability distribution function itself as minus infinity to x p x of u du, that is the relationship between the two functions. What are the properties of probability density function? Now, we have the value of probability distribution function at infinity, is actually probability of sample space and that should be equal to 1; therefore, if you evaluate this integral minus infinity to plus infinity p x of u du, it should be equal 1, that means, the area under probability density function must be equal to 1; actually, this statement is a consequence of the axiom, that the probability of sample space is 1, the actually the statement have one of the axioms.

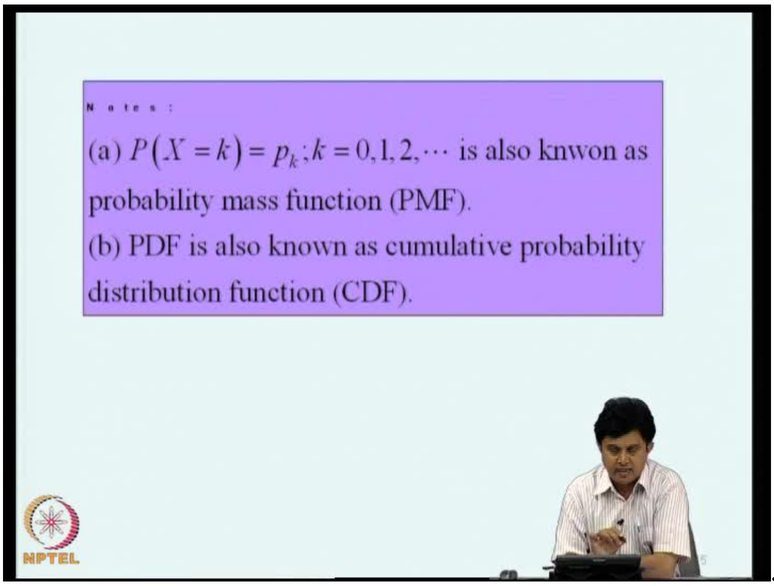
If you consider, now probability of this event X greater than equal to a and less than or equal to b, this the area under the probability density function between a and b right; so that is why, this function p x of u is called a density function. You multiply by the length of the segment, an integrate where a to b, you get the probability, the probability density function itself is not a probability, its area under the probability density function is the probability.

(Refer Slide Time: 29:51)



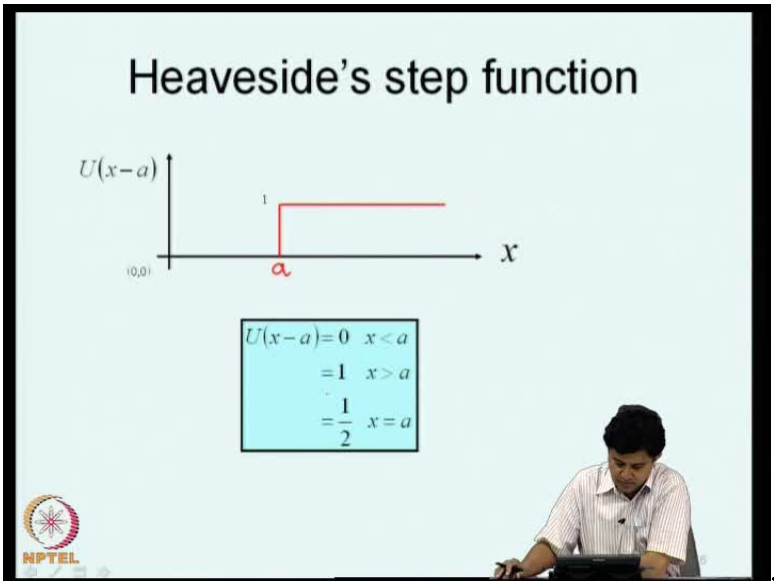
Now, this can be seen more clearly here, this is the probability density function, if you consider two point a to b, the area under this curve is a to b p x of x dx, this is the probability of a x greater than a less than or equal to b. Now, if you shrinks b to a, and consider an infinity small trap, here x to x plus dx, this area is p x of x dx, and this is the probability, of x line between x to x plus dx. Now, this probability is positive and says dx is positive p x of x must be positive or non-negative to be precise; that means, probability density function cannot be negative, because probability measured itself cannot be negative and the area under this curve must be equal to 1, both these statements are actually statements of axioms of probability.

(Refer Slide Time: 30:01)



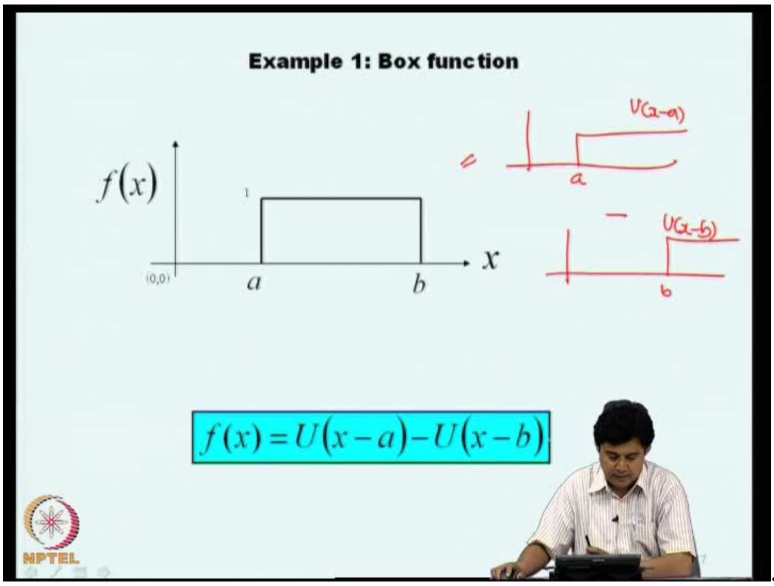
There are other terminologies that you may come across, if you read certain text books, for example, probability of X equal to k is called as is written as p k and as k is varied, and if you plot, now this function as the function of k, this is also known as probability mass function PMF. This probability distribution function is also known as cumulative probability distribution, some authors call it as CDF also, but I will strict to a notation capital PDF is probability distribution function; lower case PDF is probability density function. And I will not use these terminologies - probability mass function and cumulative distribution function.

(Refer Slide Time: 30:54)



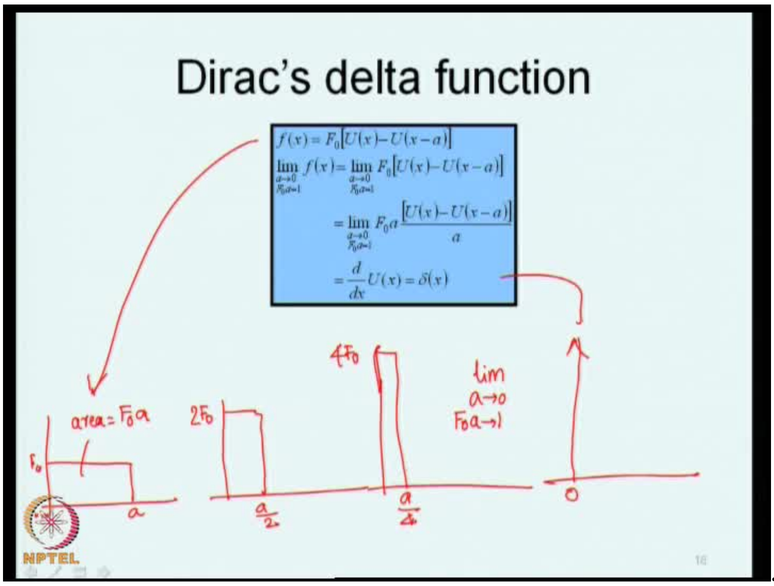
Before we proceed further, we need to now digress a bit and I need to introduce the notion of a Heave side's step function and a Dirac's delta function. Heave side's step function is shown here, U of x minus a, is set to be the Heaviside's step function. If this is at a for x less than a, this function is 0; and for x greater than a, this function is 1; and at x equal to a, we take the average of the right and left limit. And by again we call it as half that is the definition of a Heave side's step function, is also called is simply as step function.

(Refer Slide Time: 31:44)



Now, this quite a useful function to represent functions like this, for example, if you consider a box function that is f of x is 0 from 0 to a, at a there is a jump and this is the box. How can I write this in terms of step functions? We can write this function as being equal to minus b, this is nothing but U of x minus a, this is U of x minus b; so, if you detect this from this, you will get this function and you write it us U of x minus a minus U of x minus b.

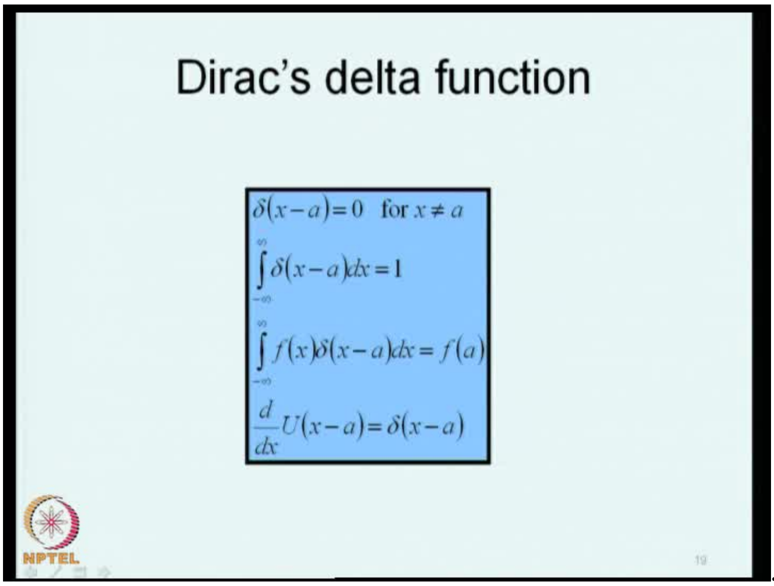
(Refer Slide Time: 32:42)



Now, what will do is, we will use this representation to describe what is known as Dirac's delta functions. So, what will do is, we will start with a function let us call this point as a, and this as f naught, and this is the origin; now, what I will do is, I will go and reducing a, but at this same time, this area, area is what? F naught into a, I will keep the area as constant equal to 1 and I will go and reducing a; so, what I will do, for instance, I will now say a becomes a by 2, this will become 2 F naught.

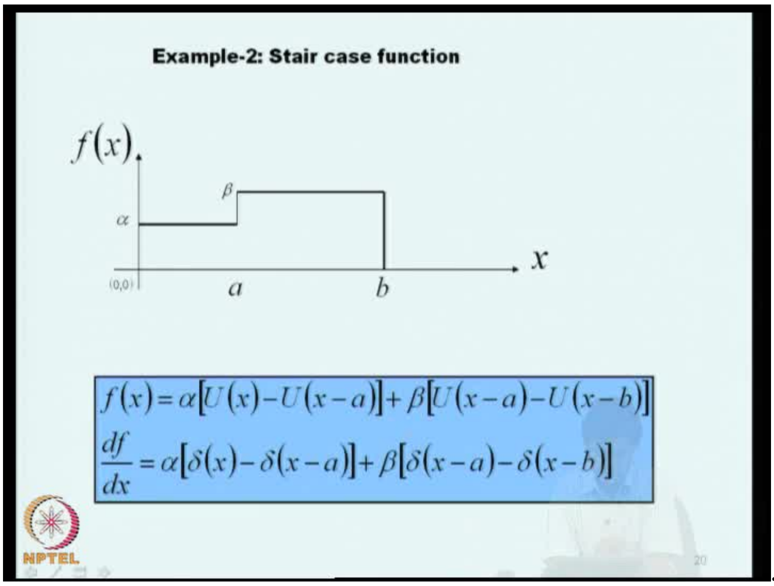
So, what is the area under this curve, 2 F naught in to 2 a by 2, it is f naught a. Now, suppose, this becomes a by 4, suppose this is a by 4, this will be 4 F naught, so the area would be still 1. So, in the limit of a going to 0, and F naught into a going to 1, what would happen to this function, it will be a kind of a spike at 0. Now, let us look at this details here, f of x is F naught into U minus x x minus a, that is this function, that is this.

(Refer Slide Time: 35:08)



Now, I am considering the limit a going to 0 and F naught a going to 1; so, I will rewrite this as F naught into a divided by a; so, as F naught into a goes to 1 and a goes to 0, you were looking at the derivative of U of x, this is U of x minus a divided by a; as a goes to 0 is nothing but du by dx and F naught a goes to 1; therefore, f of x now becomes this function and this we call it as Dirac's delta function and we represent this as a simply a spike like this. The definition of this function is as follows Dirac delta of f x minus a equal to 0; for any x which is not equal to a, but at x equal to a itself, we do not define this function; we simply say the area under this curve is 1.

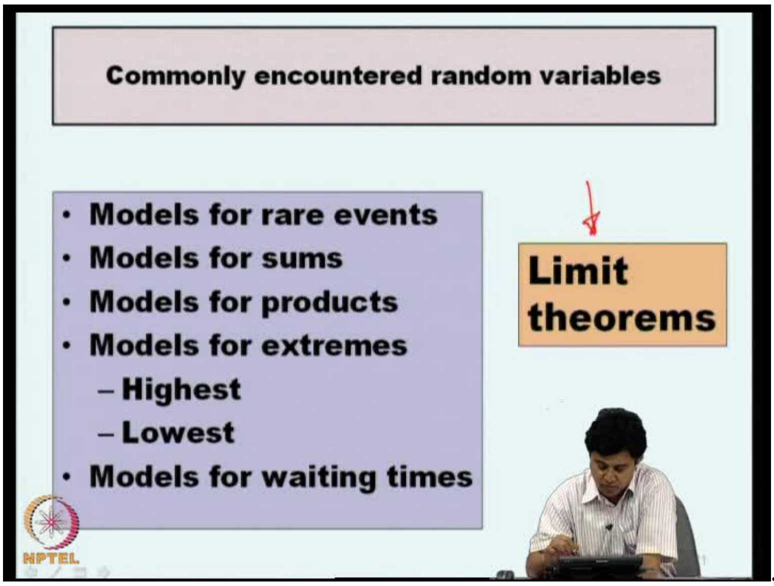
(Refer Slide Time: 35:52)



There is another property of Dirac's delta function which will need, if we take the integral of f of x into delta of x minus a dx, it will be f of a, and this property we already saw the derivative of a step function is Dirac delta function. Some of these we will be using quite often later in the course, the one more small example that will help us to understand a step function and Dirac's delta function further. So, I consider now a box which has a step here, this function can be written as alpha into U of x minus U of x minus a, that is this part, this part, plus beta in to U of x minus a minus U of x minus b, that is the second part.

So, we are able to write this type of functions in terms of step functions. Now, if we are to, if we are to need the derivate of f of x, although the slope is 0 everywhere, but at the points of discontinuity we cannot differentiate that; so, we represent the derivative in a formal way, in terms of these Dirac's delta function, suppose if we differentiate this with respect to x, this is Dirac delta function, this is direct at x minus a, this is direct delta at x minus a, direct delta x minus b.

(Refer Slide Time: 37:07)



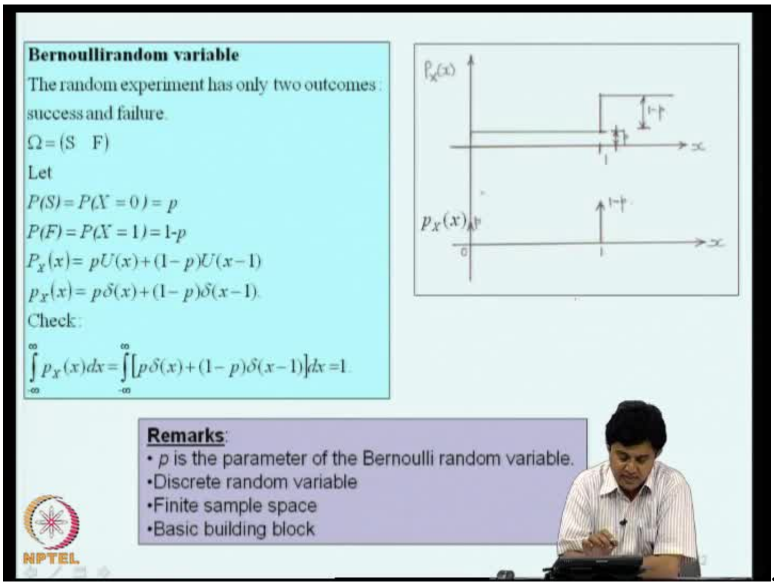
Now, why we should need, what is the use of this, I will come to this shortly. Now, I remark that the random variables help us to quantify uncertainties; now, the description of a random variable, so far, we have derived, we have introduce the notion of a probability distribution function and a probability density function. If you know probability distribution function, you can get probability density function; similarly, if you know probability density function, you can get probability distribution function. So, a complete description of a random variable is through its probability distribution function or through its probability density function.

However, you should also understand complete description of a random experiment itself is through sample space, event space and the probabilities, that triplet probability the sample space, event space and probability measures now are encapsulated in probability distribution function or probability density function. Now, just as we encounters certain types of functions, in our engineering practice or in a physical modeling a physical world like sine functions, cosine functions, exponential logarithms etcetera, they are commonly encountered in modeling, that happens, because these type of functions are obtained by solving the governing physical loss, for example, when a pendulum swings through small oscillations, the sine and cosine functions emerge naturally; in a diffusion type of model exponential functions emerge naturally. In a similar fashion, when we talk about uncertainties, there are certain kinds of uncertainties which keep occurring frequently, that would be in, there are certain types of functions, that is random variables which keep regarding frequently in modeling.

I will now considered some examples, for example, models for rare events, models for, some of you know uncertainties models for products, models for extremes, the highest and the lowest; we are interested in highest extremes of environment, which acts as input storage engineering systems, for example, floods and drop are extremes of certain climatic conditions; we are interested in strongest earth quakes and lowest strength in the structure highest wind speeds, right. So, when we talk about structural strength, we are interested what is the weakest point; when we are interested in loads, we are interested in highest points.

Similarly, we are interested in models for waiting times, it one example for a waiting time is the life of the structure itself; how much should I wait till it structure collapses, if the structure is expose to uncertain effects and it has uncertain properties its life also is a random is uncertain. So, while introducing random variables, it will be quite useful; if you look at random variables from a certain physical perspective like this.

(Refer Slide Time: 40:55)

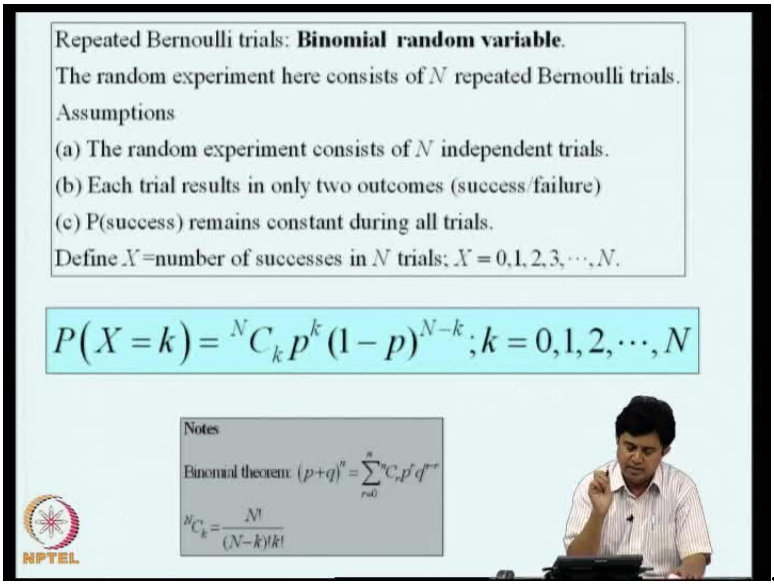


This is another important property of random variables that we should understand; these are associated with limit theorems. So, these are very important results in probability theory, I will come to that due course; so, we will be now going through some of these in sequence. The first random variable that will be looking for that will be studying is known as Bernoulli random variable. In a Bernoulli random variable, the underlying random experiment has only two outcomes, call it as success or failure. It could be for example testing a cube in a laboratory and see whether the compressed strength exceeds a threshold value or not; there only two outcomes here, toss a coin you get head or tail, you appear for an exam, you pass or fail, there only two outcomes.

We call probability of success, we plot the random variable, success we map to 0, failure we map to 1, and probability of success we call it as P, and therefore probability of failure must be 1 minus p. So, what is the probability distribution function, here this is the origin (0, 0), this is 1; so, the claim here is P and the claim here is 1 minus p; so, this is the probability distribution function. So, what can you say about this probability distribution function with corresponds to a discrete random variables; so, the sample space here has two outcomes, so it is the finite sample space and this is a discrete random variable, and if you want to construct its probability density function you have use the representations using step functions; so, you get one Dirac delta function at the origin which is p and other Dirac delta function here, where the climb is 1 minus p and that is what appears.

So, this is the probability density function, this is the probability distribution function. So, if you look at these plots, all the underlying features of your modeling of random experiment or encapsulated here, it is in this sense that I say that, this helps in encapsulation. Now, the probability distribution function is p in to U of x plus 1 minus p in to U of x minus 1; now, the density function is the derivative of this p in to direct delta plus 1 minus p in to direct delta x minus 1, is the area under the curve equal to 1, of this, you find the integral this is p delta x plus 1 minus p delta of x minus 1 dx; now, p into delta x area under that curve is p 1 minus p, area under this curve is 1 minus p, so p plus 1 minus p is 1, so all axioms of probabilities are respected; so, this is the how we describe a simple random variable like Bernoulli random variable.

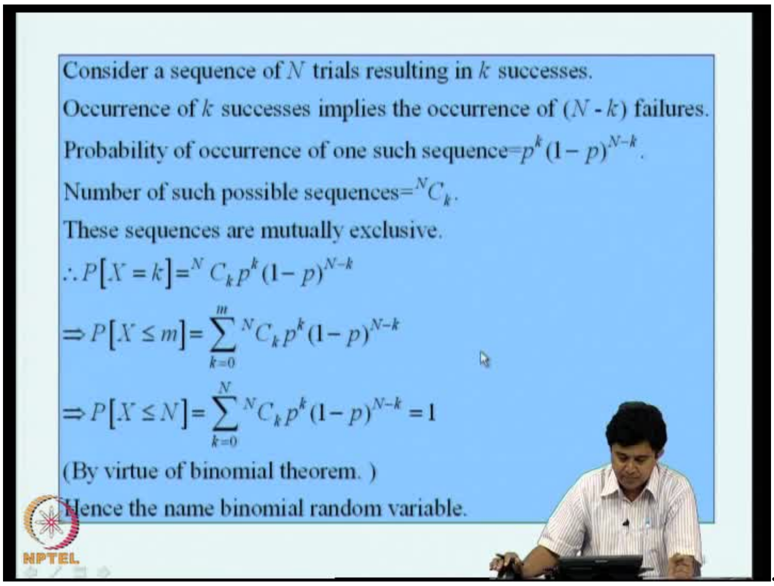
(Refer Slide Time: 43:57)



This Bernoulli random variable is an important random variable; in the sense, it forms the building block for constructing many other random variables; so, we call it as binomial random variable. The sample space here consist of repeated Bernoulli trials, you toss a coin for n number of times, and ask the question, how many heads you have to obtain, that means, the random experiment here consist of N repeated Bernoulli trials. We make certain assumptions; here the random experiment consists of N independent trials, each trial results in only two outcomes success or failure that is what I mean, by saying that the trial is the Bernoulli trial. Probability of success remains constant during all trials; suppose, we are tossing a coin at every toss of the coin the probability of getting head should remain the same. Then, we define X as a number of successes in N trials; so, how many successes are possible, you may not succeed at all or you made succeed all the time; so, it is 0, 1 2 3 N.

The probability of X equal to k, we shortly show this can be shown as N C k p k 1 minus P N minus to the power of N minus k, where k take values from 0, 1, to N; now, why this random variable is called binomial random variable? The reason is, if you have to find out probability of sample space; that means, you have to sum this probability from k equal to 0 to capital N, in doing, so you utilize the binomial theorem which says that p plus q to the power of n, is given by this. We utilize this in showing that this description obeys one of the axioms of probability namely, probability of sample space is 1; therefore, the name binomial random variable.

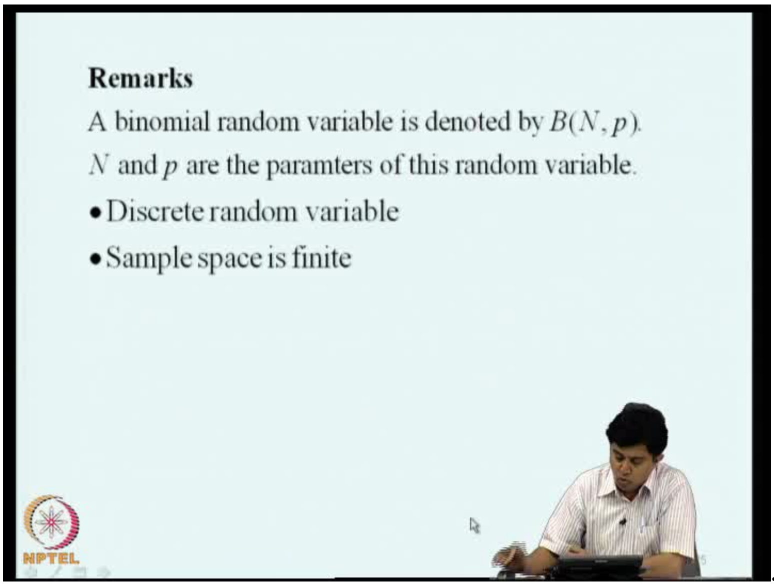
(Refer Slide Time: 45:54)



How do we get this probability of X equal to k, to be this, we will see now in a minute. Now to derive this description, we consider a sequence of N trials resulting in k successes, I have tossed the coins in 30 times and I have got 8 heads; so, I am trying to describe that. Now, occurrence of k successes implies the occurrence of N minus k failures, if I got 8 heads, I should have 22 tails. Probability of occurrence of 1, such sequence is p to the power of k, k successes and N minus k failures and all these are independent; therefore, I multiply these probabilities p to the power k 1 minus p N to the power of minus k.

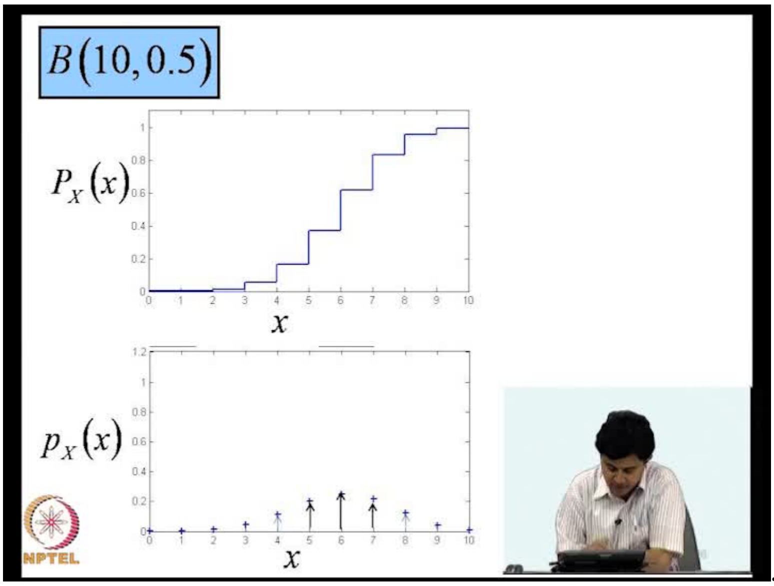
How many such sequences are possible, how many ways I can draw k objects from a pile of N objects N C k and these sequences are mutually exclusive, if one sequence occurs, the other sequence cannot occur; so, the intersection is the null set. Therefore, if we now look at probability of X equal to k, I have to, I can use now the third axiom of probability and add all these sum; so, I get N C k p to the power of k 1 minus p N minus k, this is how the probability mass function, as we call for this quantity is derive, for a binomial random variable. Now, what is the probability distribution function? It is the probability of X less than or equal to m k, running from 0 to m this, and probability of X less than or equal to N is actually this summation, and you can shows this to be 1, by virtue of binomial theorems. As I said this is why this random variable is called a binomial random variable.

(Refer Slide Time: 47:37)



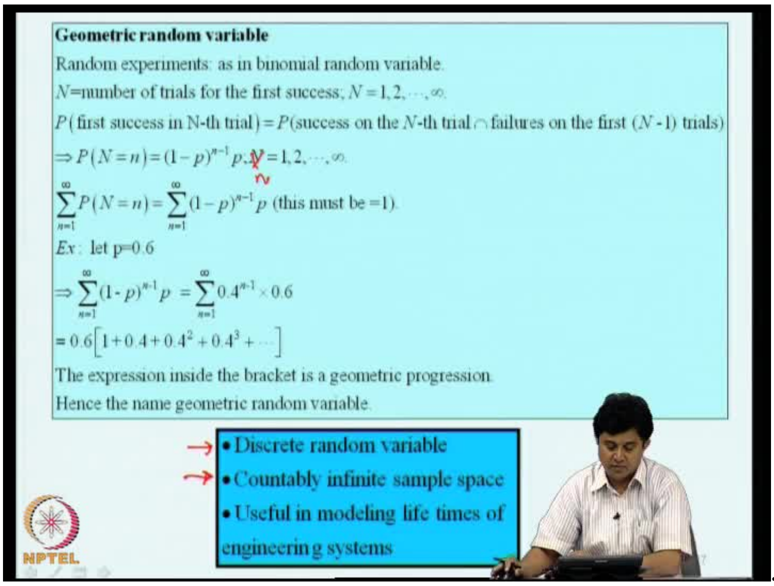
Now, there are two parameters in this description of random variables, this random variable, namely, the number of trials and the probability of success on any given trial. So, we see, we say that, a binomial random variable has two parameters capital N and p. This is again a discrete random variable; the sample space is finite, because there is only N plus 1 outcome, 0, 1, 2, 3, 4, and up to capital N.

(Refer Slide Time: 48:06)



Now, a plot of a binomial probability distribution and density function of a binomial random variable, where capital N is 10, you can plot that function and it appears as the staircase; this is the probability distribution function. And associated with this, there are spikes corresponding to this jumps and this is the probability density function.

(Refer Slide Time: 48:34)



Now, based on binomial random variable, we can define what is known as a geometric random variable. Now, the random experiment here is, as in binomial random variable; that means, you will go and tossing coin N number of times and see how many heads you get, but the observation will make here is, this random variable is how many trials are needed for the first success, that is the random variable; here how many times you need to appear for exam, before you pass or before you know cross a certain threshold.

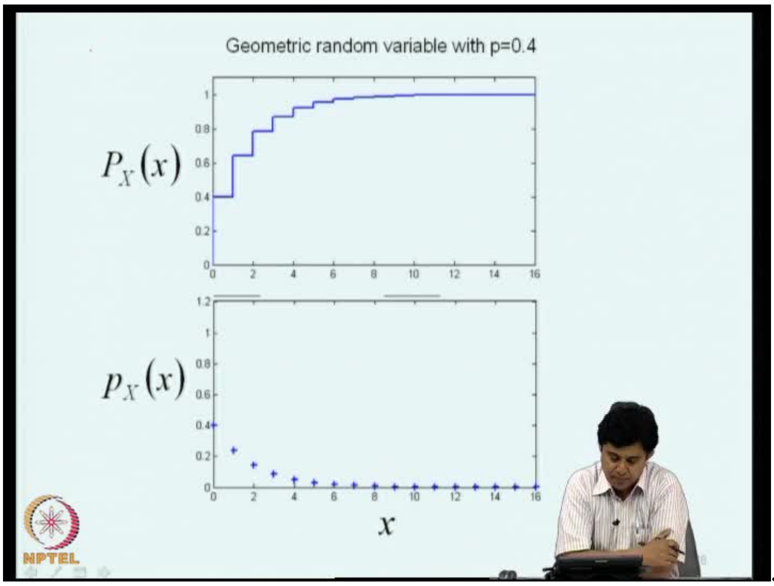
Now, how do you construct the, this probability; now, let us consider that the successes occurred in nth trial, the first success probability of first success in nth trial is, if you have go to head on thirtieth trial, you would have tails on first 29 trials, because thirtieth trail is the first success; so, probability of success on nth trial intersection failures on the first N minus 1 trials, if you moment you get a head you stop the experiment, right. So, here what is therefore probability of N equal to n, it is 1 minus p to the power of N minus 1 in to p.

Now, what happens if you sum this states are here 1, 2, 3 up to infinity; so, what can you say about this random variable, this is the discrete random variable and it has count ably infinite sample space, you will, you may never succeed; so, you may have to go and tossing a coin, the coin is so bad that will never get a head; so, the number of trials can extend all the up to infinity so it has count ably infinite sample space. Now, is the axiom of, one of the axiom probability of sample space is equal to 1 or not, how do you check? You have to sum this probability from N equal to 1 to infinity, if you do that, you get this and for the axiom to be obeyed, this must be equal to 1; so, to be illustrate as let us consider p equal to 0.6.

So what is this, summation N equal to 1 to infinity 1 minus p N minus 1 in to p, this is this and if you expand and write this is 0.6 1 plus 0.4 etcetera. So the number in the bracket is the geometric progression and you will see that, if you use the result of sum of a geometric progression, you will see that this area is 1. So, in the name geometric random variable, now arises from the property that we are to show that probability of sample space is 1; we are using a property of a geometric progression, so that is why this in the term geometric random variable.

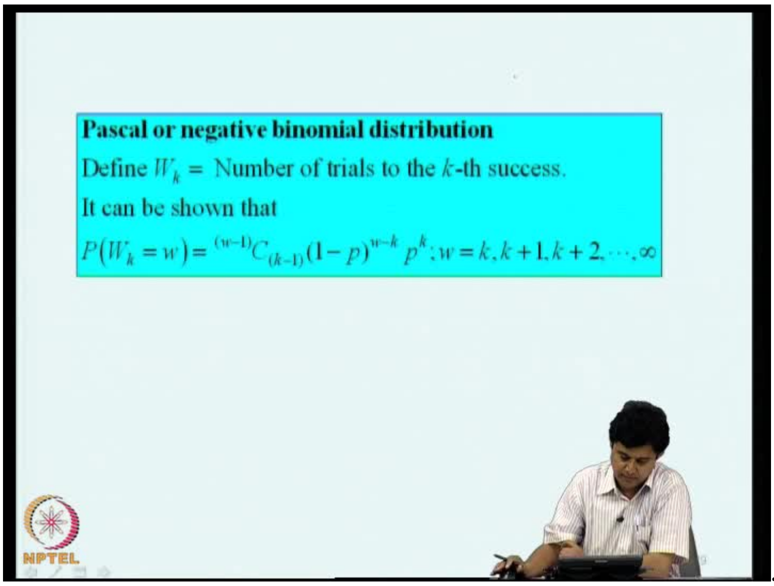
Now, this geometric random variable is a very useful modeling tool, if you are trying to model lifetimes, suppose at a point on the structure, we say the structure as failed, if stress exceeds a threshold, some threshold; so, I every year or every time unit, I monitor how many years I have to wait till the response of the structure exceeds a threshold. So, these are typical question that we need to ask and you can immediately see that, this geometric random variable has relevance in answering such questions.

(Refer Slide Time: 52:18)



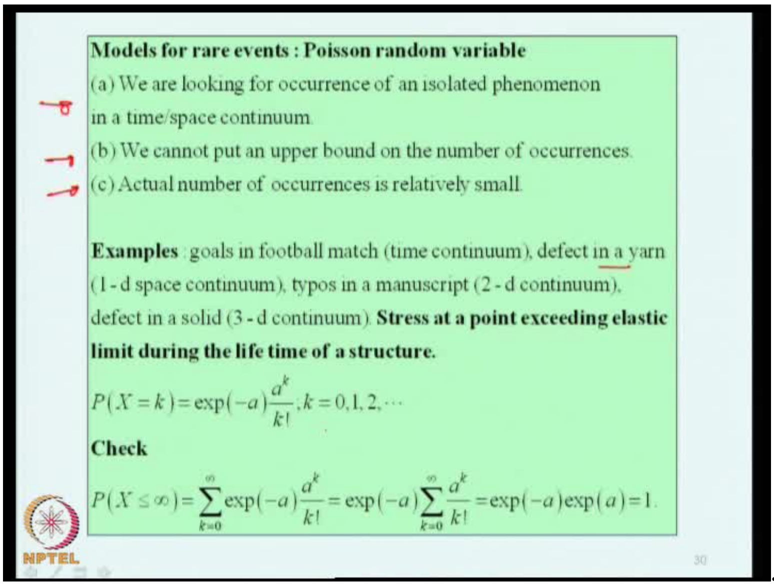
This is the one plot of the probability distribution of geometric random variable, where the P, that is the probability of success on any trial is taken as 0.4 and here, again you will see that, this is the discrete random variable, and it has count ably infinite sample space; so, there are infinite number of this small steps and eventually this distribution function reaches 1. The derivative of this, again it does not actually exist, but in terms of direct delta function, we can represent and we get a spike at these places, where there are jumps. So, this is the probability density function, this is the probability distribution function.

(Refer Slide Time: 53:02)



There is the another generalization of geometric random variable; we define W k as number of trial to the kth success; in the geometric random variable, we waited for first success. Now, here we are waiting for kth success, this could be a relevant, for example, if you are modeling fatigue phenomenon and so on and so forth. Here, you can show that this W k the probability of W k equal to W, is given by this expression; this can be shown using properties of binomial random variables.

(Refer Slide Time: 53:48)



Next, we move want to models for rare events; here, we come across what are known as Poisson random variables, the story here is as follows. Here, we are looking for occurrence of an isolated phenomenon in time or space continuum. We have a simple example in mind, suppose you are looking at number of goals code in a football match, so you are looking at isolated events in time continuum. If you are looking for a defect in a yarn, you take a yarn and you inspect for defects, you are looking for isolated phenomena in space continuum, one-dimensional space continuum.

Similarly, if you are looking for typographical errors in a manuscript, you are looking for isolated events in a two-dimensional continuum. If you are looking for defect in a solid, you are looking for an isolated event in a three-dimensional continuum. If you are looking at time history of a stress at a point, one of the components of stress at a point in time and you are looking at the event, that it exceeds the elastic limit, during the lifetime of the structure, you are looking at isolated phenomena in time.

Now, if you look at for example, goals in a football match, we cannot put an upper bound on the number of occurrences, we cannot say in a football match goals more than, so many are not possible, any number is possible, but in reality, the actual number of occurrences is relatively small; you would not hear of a football match, where 800 goals are scored, if it is 4 goals, 5 goals are ok. If these conditions are satisfied, this counting random variable, where we count the number of this isolated phenomena follows what is known as a Poisson distribution. This is the random variable, it is the integer valued random variable, it has count ably infinite sample space and the probability mass function probability X equal to k, it can be shown that it is given by exponential of minus a a to the power of k divided by k factorial, where k runs from 0 1 2 up to infinity.

We can show that the some of this, that means, again at whenever we introduce a probability distribution function, we need to check whether the axiom of probabilities obeyed or not; so, if you sum this quantity from k equal to 0 to infinity, you can show that this sum becomes 1.

So, we conclude this lecture at this point, we continue this in the next lecture.