

Probability Methods in Civil Engineering
Prof. RajibMaity
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

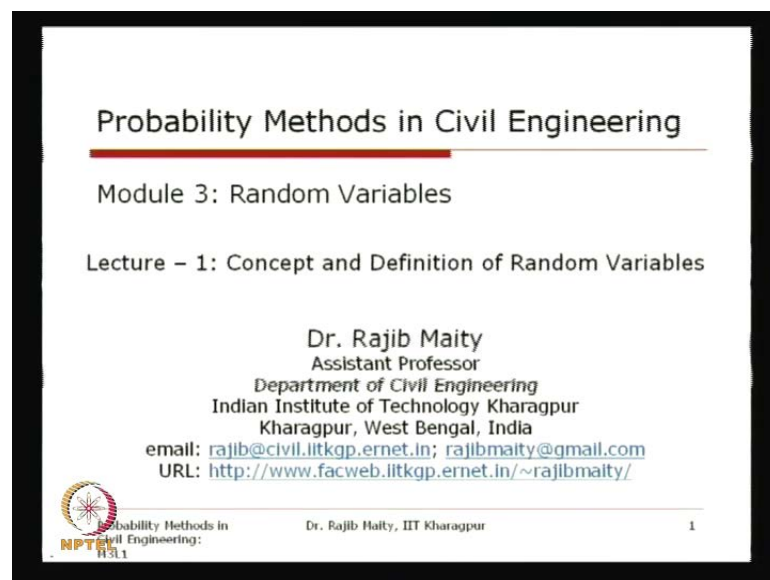
Module No.# 03

Lecture No. #06

Concept and definition of Random variables

Hello there and welcome to the course probability methods in civil engineering. Today, we are starting a new module, module 3.

(Refer Slide Time: 00:28)




Probability Methods in Civil Engineering

Module 3: Random Variables

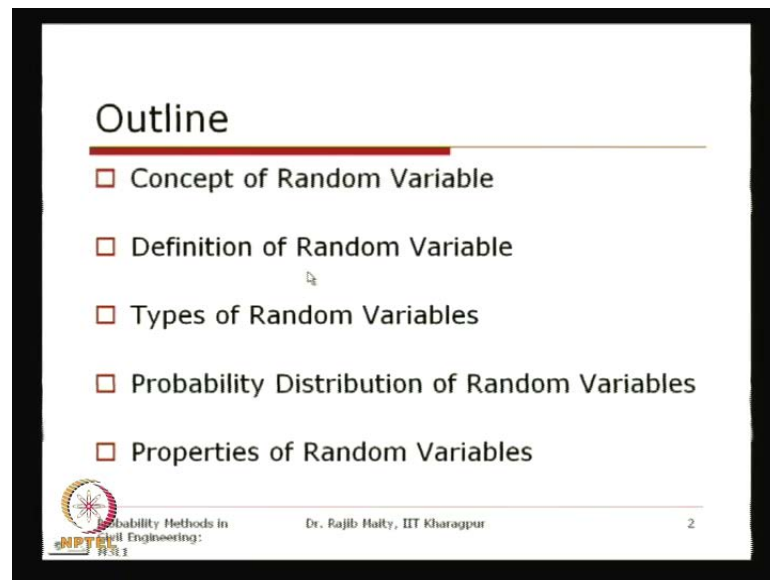
Lecture – 1: Concept and Definition of Random Variables

Dr. Rajib Maity
Assistant Professor
Department of Civil Engineering
Indian Institute of Technology Kharagpur
Kharagpur, West Bengal, India
email: rajib@civil.iitkgp.ernet.in; rajibmaity@gmail.com
URL: <http://www.facweb.iitkgp.ernet.in/~rajibmaity/>

 Probability Methods in Civil Engineering: Dr. Rajib Maity, IIT Kharagpur 1

In this module, we learn about random variables and there will be couple of lectures and in this first lecture, we learn the concept of random variables and their definitions.

(Refer Slide Time: 00:45)



So our outline for today's lecture is arranged broadly like this. First, we will discuss about concept of random variable and then, we will understand the definition of random variable. Then, we will discuss different types of random variables and their probability distributions.

This random variables are important in the sense, that this is the, this concept is important to understand how it is used in the probability theory and this probability theory for this random variable is generally, is is generally assessed through the through their distributions. So, this is known as the probability distribution of random variables. After that, we will learn different properties of random variables.

(Refer Slide Time: 01:40)

Concept of Random Variables

Concept of Random Variable (RV) is most important to the probability theory and its applications

Functional correspondence

A Random Variable (RV) is not a variable!

Probability Methods in Civil Engineering: NPTEL 11.1

Dr. Rajib Halty, IIT Kharagpur

3

So, as just what we are telling, that this concept of this random variable, which is abbreviated as rv, is most important to the probability theory and its application. Now, if we see that, how what actually this random variable mean, if this shaded area, what you see here is the **is the** sample space, then this sample space consist of all feasible outcome of one experiment.

Now, this feasible outcome of this experiment can be continuous or can be discrete points from through this sample space. Now, from the sample space, we have to, for the mathematical analysis, we have to map the outcome of the experiment, of one experiment to **theto** some number, according to our convenience.

Now, to map this, each and every output of a random experiment to this real line, is generally through a functional correspondence and this correspondence, this functional correspondence is one random variable.

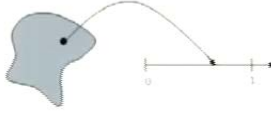
So, what is important and what is, what we want to understand from this slide is that, this random variable, even though this variable is shown here in the name of the random variable, is not a variable. So, which is generally, what you want to say is a functional correspondence or is a function. That function generally map between the sample space of one experiment to the **to the** real line.

Now, this mapping is done based on our convenience and that we will see in a minute.

(Refer Slide Time: 03:36)

Concept of Random Variables

- A Random Variable is a function denoted by X or $X(\cdot)$, that map each points (outcomes of an experiment) over a sample space to a numerical value on the real line.
- Notations:
Generally a Random Variable is denoted as uppercase letter and its specific values are denoted as lowercase letters. Thus,



Random variable $\xrightarrow{Y(x)}$ Specific Value

Probability Methods in
Civil Engineering:
IIT Kharagpur

Dr. Rajib Halty, IIT Kharagpur

4

Thus, if we say the definition, if you say, then we will say that a random variable is a function denoted by x or x some specific number, it can be denoted by any **any** letter. Generally, these things are denoted by this small lower case of this **of this** random variable. These random variables are generally expressed in terms of the uppercase letters. So, this random variable x is a function that maps each point, that is an outcome of an experiment over a sample space to the, to a numerical value on the real line. What just we have shown in this figure, in the last slide.

Now, before we discuss some of the, some of the further concept of this random variable, so there are different notations are being used to denote this random variable and the specific value. For this lecture or you will see in some of the classical text book, that this notation will be followed for this **for this** course. Generally, this random variable are denoted by the uppercase letters and its specific values are denoted as lowercase letters. Thus, in this notation, this uppercase letter is indicating the random variable, which is denoted by this x and any specific value to this random variable is denoted by this small x , which is denoted, which is shown here. So, this is the specific value of the random variable x .

(Refer Slide Time: 05:20)

Concept of Random Variables

□ Examples:

a) Number of rainy days in a month
denoted by D . Some value, such as $d=5$ indicates 5 rainy days are observed in a particular month, expressed by $D(d=5)$

b) Number of road accidents over a particular stretch of a road in a year
denoted by X . Some value, such as $x=10$ indicates 10 road accidents over that stretch of road has occurred in a particular year, expressed by $X(x=10)$

c) Strength of concrete, C
denoted by C . Some value, such as $c=25 \text{ N/mm}^2$ indicates 25 N/mm^2 strength was observed for particular sample of concrete, expressed by $C(c=5)$

Reliability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 5

Now, we will see some **some** example of this random variable which we just discussed. Suppose that, number of rainy days in a month; so, if we can, you can say that this is a random variable and we do not know. So, this value that number of rainy days in a month can vary and if we denote this random variable as letter capital d and some value say d equals to 5. So, this indicates that 5 rainy days are observed in a particular month. Now, which is, which can be expressed by d in parenthesis d equals to 5.

Second example, say the number of road accident over a particular stretch of the road in a year. Suppose, this is denoted by x and some specific values, such as say x equals to 10, indicates that there are 10 road accidents over that stretch of the road has occurred in a particular year. Now, this value 10 or this value 5 can change in the successive **successive** time step or in the success; for in this case, in the successive year, that can change. So, this number of road accidents is my **is my** random variable, which for a particular year it takes a specific value x equals to 10, which is denoted by this.

The last one, say the strength of concrete. If, I designate it as c , so this is denoted by c and some specific value such as c equals to 25 Newton per millimetre square. It indicates that 25 Newton per millimetre square strength was observed for a particular sample of concrete and which is expressed by c in parenthesis c equals to 5. C is the, **sorry** this shall be c equals 25; **sorry** for this mistake this should be 25.

(Refer Slide Time: 07:25)

Definition of Random Variable

- A random variable X is a process to assign a number $X(\zeta)$ to every outcome ζ of a random experiment
- The resulting function must satisfy following two conditions:
 - The set $\{X \leq x\}$ is an event for every x
 - The probabilities of the events $\{X = +\infty\}$ and $\{X = -\infty\}$ are equal to zero

Papoulis and Pillai, 2002

Probability Methods in
Civil Engineering:
NPTEL

Dr. Rajib Halty, IIT Kharagpur

6

So, a little bit formal definition of random variable. If we see, that a random variable x is a process to assign a number, which is a real number x_i to every outcome ω_i of a random experiment. So, if we just take one, take that thing, so suppose that we are taking that example of that of tossing a dice, then we know that whatever the numbers that generally comes, if it comes a 6 on the toss, then we see, we say that the outcome that is coming, it is my 6. Similarly, six different surfaces can have six different numbers and that can be happened.

(Refer Slide Time: 07:51)

The diagram shows a 3D cube representing a random experiment. The faces are labeled B (top), R (right), and G (left). To the right, a mapping for a random variable X is shown:

Outcome	Value of X
B	10
R	20
G	30
⋮	⋮
⋮	⋮
⋮	60

© CET
I.I.T. KGP

NPTEL

Now, see that in this experiment, if I just change this one, so this is quite straight forward that, whatever number of dots that we are usually usually used to know. Suppose that, instead of having this kind of dice, if I just have a have a dice which is having different colours. Suppose that, this is having a blue colour, this side is having may say red colour, this side is having green colour and similarly six different faces having six different colours. Then, what I am trying to do is that, this outcome of this is a top surface, the colour of the top surface, if I just 1 to 1 to say that this is my outcome of this of this experiment, then what I can do is that I can take different colour code and I can assign to some number, not necessarily 1 to 6, it can be of any number.

Say that b, I am giving some number 10, r giving some number 20, green giving some number 30 and this way say up to 60 this number is given. So, these mapping from this outcome to some number, this mapping is your random variable. So, we will come after this, how to assign the probabilities. Now, this is very straight forward. So each and every outcome, if this dice is fair, then the outcome of each and every of this outcome will be equal; so everything will be 1 by 6.

So, that is again another process. How we will assign the probability to that particular random variable? If it is x , so that will be, how we assign the probability. Now, when we are defining this random variable itself, so this random variable is nothing but, is a process to assign a number to each outcome of that random experiment. This is exactly what we are calling it as a formal definition of this random variable is that, a random variable x is a process to assign a number x_i to every outcome ω_i of a random experiment. The resulting function must satisfy the following two conditions.

Even though, we are saying that this is a this is a function, so this function of mapping from this outcome, experimental outcome to some numbers should follow two, such two conditions. The first condition is the set x less than equals to small x . Now, you see here, this is the capital letter which is your random variable and this is the specific value. Now, this random variable, below some specific value is an event for every x . So, whatever the x you take, so this should constitute an event for that random variable.

For that, for this x , now other things of probabilities of the events, that x equals to plus infinity plus infinity and x equals to minus infinity are equal to 0. Why these things we are

just saying is that, even though means mathematically, this x can take any value on the real line but still, we say that so the extreme, that is plus infinity and minus infinity, this probabilities if I take this probabilities, this should be **should be** equals to 0.

(Refer Slide Time: 11:40)

Definition...contd.

- A complex random variable ($\mathbf{z}=\mathbf{x}+j\mathbf{y}$) is only defined by joint distribution of random variables \mathbf{x} and \mathbf{y} , since the inequality $\mathbf{x}+j\mathbf{y}\leq\mathbf{x}+j\mathbf{y}$ does not exist
- If not stated all random variables are real

Probability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 7

A complex random variable, so just we will to complete this definition part, if we do not say this one, it will not be completed. So, this is, there is another type, this is the real random variable which just we discussed, is now it can be the complex random variable as well, which is z equals to x plus j y .

So, this random variable is only defined by their joint distribution. So, if I want to define the complex random variable z , this should be defined by their joint distribution. This joint distribution will be discussed later in this **in this in this** course. So, but for the time being, you can say that this joint distribution is generally obtained from that distribution, of this x and y both. So this both, this distribution should be taken care to find the joint distribution to define the distribution of the complex random variable, of this, of the random variable, of individual x and y . Since, this inequality because, so this kind of inequality, say this of the **of the** random variable and this is the less than equal to that number does not exist. But, for this course, if it is not stated specifically, all the random variables are real. We should follow this norm.

(Refer Slide Time: 13:15)

Types of Random Variables (RVs)

1. Discrete Random Variable
2. Continuous Random Variable
3. Mixed Random Variable

■ Discrete Random Variable

- If the possible set of values that a Random Variable can define is *finite*, it is a Discrete Random Variable
- Probabilities are defined for specific values, which are greater than or equal to zero.
- Summation of probabilities for all possible values are equal to 1

Probability Methods in Civil Engineering: IIT Kharagpur 8

Now, with this concept, we will just see different types of random variables. There are basically two types of random variables; one is the discrete random variable and the other one is continuous random variable. But, there is, there are some times, in particular in the civil engineering application, we have seen there are some distributions which can be treated as the mixed random variable as well. So, we will discuss these different types of random variable one after another now. The first we take this discrete random variable. If the possible set of values that a random variable can define is finite, it is a discrete random variable.

Now, so just to tell this thing here, that suppose that this is my **this is my** sample space. Now, if I say that this random variable is defined for some specific **specific** value and rest of the part it is not defined, so it takes some specific **some specific** value. So, those values only it can **it can** define. So, in that case, what we say that **this is** this is a discrete case. So, in the in the previous example, if we just see here, that it can define that some specific outcome of this **of this** random variable. It cannot, if it is not that continuous, that a range cannot be **cannot be** defined **defined** here. So, if it is **if it is** dots, then we can say that it is defined for the outcome 1, defined of the outcome 2. But, it is not defined for the outcome anything any number in between this two.

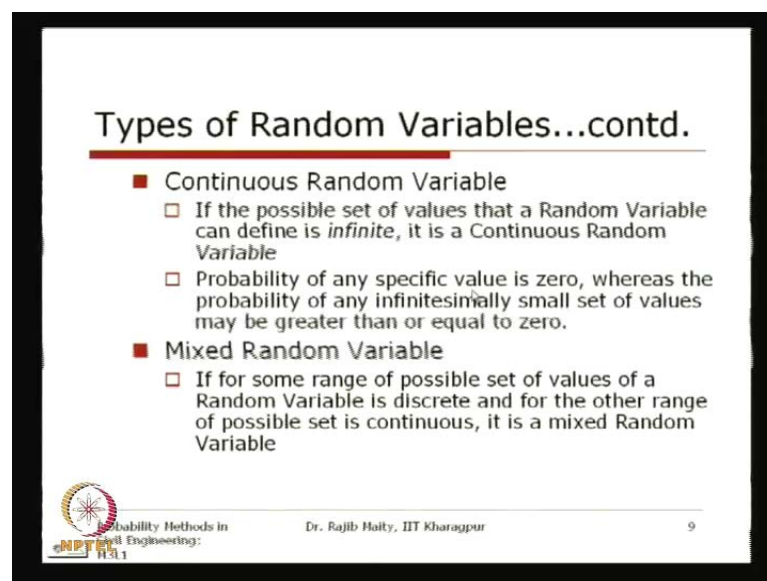
Similarly, if we just see from this remote sensing concept, that colour code, then we can say that it is defined for the blue for this coloured surface, it is defined for the blue, it is

defined for the red but, in between mixed. I think you know this colour code. This stands for this blue, if it is the colour are r g b are the primary colours and this 1 00 is for the red. Now, you can just change this number to some other to get some in between colour. So, that is not defined there. So what we want to say for this discrete is the, these random variables are described only for some specific outcome of this experiment. It is not defined for this all. So, those **those** random variables which outcome who can map, who can define the outcome of the **of the** random experiment for some specific values, those are known as the discrete random variable.

Thus, if the possible set of values that a random variable can define is finite, it is discrete random variable. The probabilities are defined for the specific values which are greater than or equal to 0. So, this we will see in detail which are greater than or equal to 0. But, what we are trying to stress here is that, the probabilities are defined for the specific value. So, it is just like that, for that point this probabilities are defined for the adjoining point, this is not defined.

Summation of all the **all the** probabilities, all the possible values are equals to 1. These two things are basically coming from the actions of the probability theory that we discussed in the previous classes.

(Refer Slide Time: 16:45)



Types of Random Variables...contd.

- **Continuous Random Variable**
 - If the possible set of values that a Random Variable can define is *infinite*, it is a Continuous Random Variable
 - Probability of any specific value is zero, whereas the probability of any infinitesimally small set of values may be greater than or equal to zero.
- **Mixed Random Variable**
 - If for some range of possible set of values of a Random Variable is discrete and for the other range of possible set is continuous, it is a mixed Random Variable

Probability Methods in
Civil Engineering: IIT Kharagpur

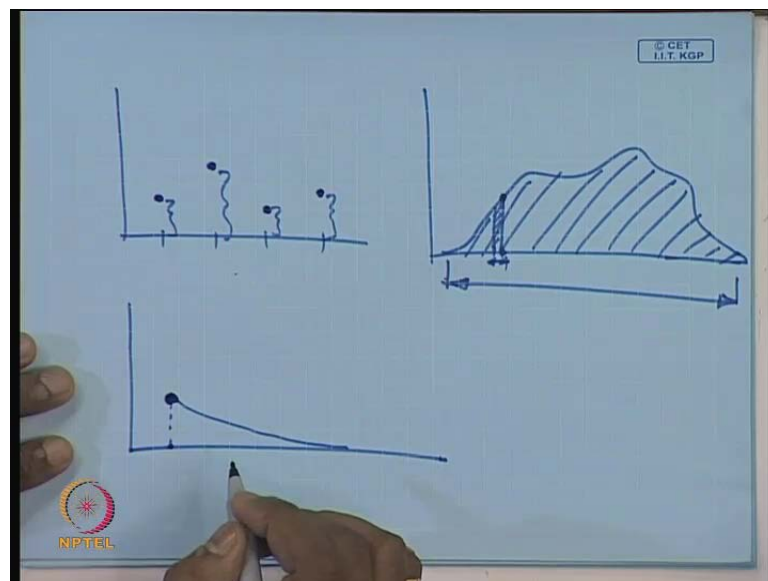
Dr. Rajib Halty, IIT Kharagpur

9

The next type is continuous random variable. So, on the other end, in contrast to this, to this discrete random variable, if the possible set of the values that a random variable can define is infinite, it is continuous random variable.

So over a zone, it is over a zone, over a range this random variable is, if it is also defined, then this is known as the continuous random variable. The probability of any specific value is 0. Now, this is coming from this density point of view. Now, once we are saying that this probability of any specific value, if I just take, then this specific value probability is 0. This is again coming from the actions of this probability. But, even though we say that it is exactly, specific value it is 0 but, whatever small set of the values if we think, that can be greater than or equal to 0. So, whatever the small range, if we **if** **we** consider, then it will be 0. I think this will be more clear from this pictorial view here.

(Refer Slide Time: 18:01)



So, suppose what we are trying to say that, if there are some specific outcome and those probabilities are assigned like this. Whatever may be the probabilities, if it is not these are, these are like the concentrated mass for that specific outcome.

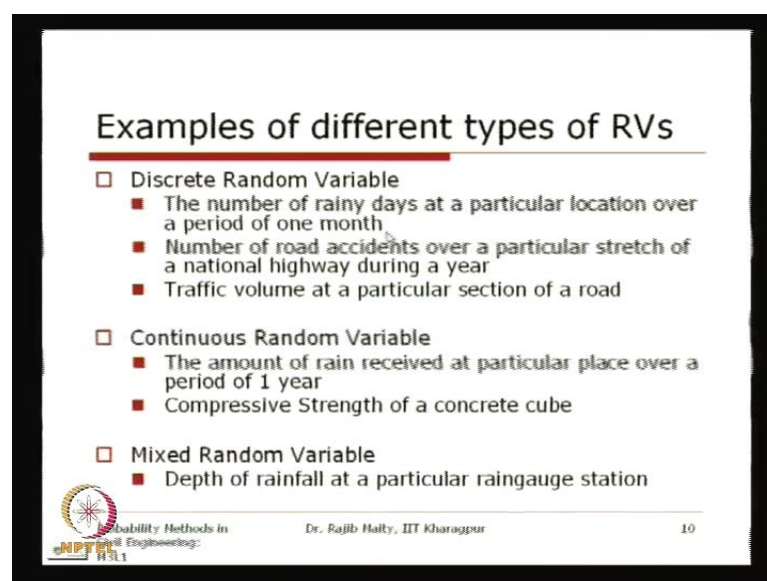
So, this is this is your discrete random variable. Now, for this continuous random variable, what is defined is that it looks like this. So over this zone, from this to this zone, for any specific value, this is nothing but, the density of this one. Now following the actions of this probability, if we just adapt with this axis, is your **is your** probability. As if, I adapt all this values to what we just discussed, this will be equal to 1. But, in this case, when we are

talking about, this is one continuous distribution then, what happens for a specific point, the probability is 0 because this is, basically, a density. Now, whatever the small area that we will consider, then this area is giving you the probability for this small zone. Now, for the full zone, if I just take then the total probability, that is, total, this is your feasible range. So the total probability in this feasible range should be equal to 1. That comes from the actions of this probability theory.

Now, again here, the next type what comes is the mixed distribution. So mixed distribution, sometimes what happens for specific value, it is defined some probability. Some probabilities concentrated here as a mass and for the rest of some range it can be continuous and goes up to some level or up to the end of this or up to the infinity. Now, this for this particular point, this probability mass is concentrated here and for the rest of the region, it is having some continuous things.

In such cases, we say that these kind of random variables are mixed random variable. So, this is the third thing, what we are seen that. So, mixed random variable says, if for some range of possible set of the values of a random variable is discrete and for the other range of possible set it is continuous, it is a mixed random variable. Now, with these three type of random variables, some example in the civil engineering context, we will see now.

(Refer Slide Time: 20:40)



Examples of different types of RVs

- Discrete Random Variable
 - The number of rainy days at a particular location over a period of one month
 - Number of road accidents over a particular stretch of a national highway during a year
 - Traffic volume at a particular section of a road
- Continuous Random Variable
 - The amount of rain received at particular place over a period of 1 year
 - Compressive Strength of a concrete cube
- Mixed Random Variable
 - Depth of rainfall at a particular raingauge station

Probability Methods in
Civil Engineering

Dr. Rajib Halty, IIT Kharagpur

10

The first is that, we discuss is the discrete random variable. Say that, this number of rainy days. Now number of rainy days at a particular location over a period of one month say, so it can take, either the specially the integer value, either 1 day or 2 days 3 days, in this way up to 30 or 31 days, whatever may be the case.

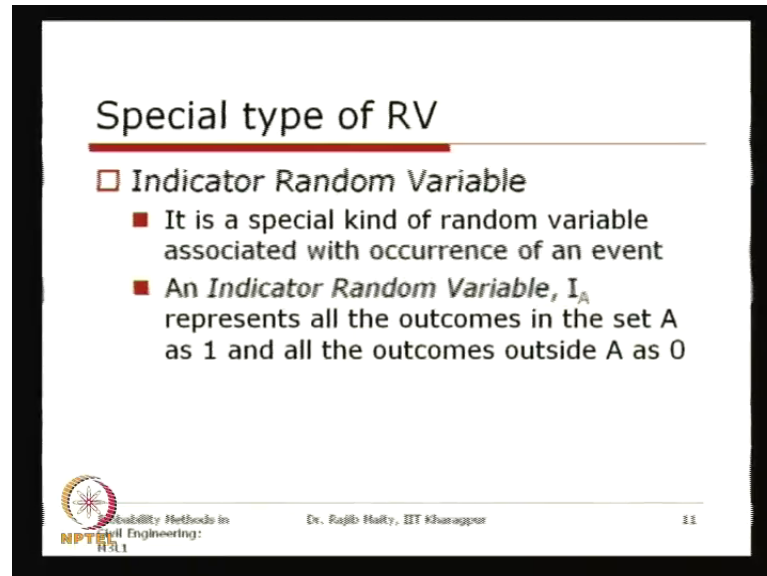
But, it can never take any value in between two integers. So, as this **so as this** can take only some specific values, so this is an example of discrete random variable. Second is a number of road accidents over a particular stretch of a national highway during a year. So again, these road accidents can be either 1 2 3 4 in this way up to infinity but, it can never take any value in between two integers. So, this also one example of discrete random variable. Third is the traffic volume at particular location of a road. So, this is the number of vehicle that is crossing that particular section of the road. Again, this can take only the integer values for that one. So, this is also one example of the **example of the** discrete random variable.

Now, the example of continuous random variable, on the other end, that amount of rain received at particular place over a period of one year. So, annual rainfall for that session, for the, say, it is the number of rainy days, which is a discrete. Now, if I just say that total depth of the rain received at that particular point that can vary from 0 to any possible any possible number. So, this is one example of the continuous random variable. Compressive strength of a concrete cube; now, there are different ways, there are different criteria, maintaining different design criteria, if we prepare the concrete cube and then, we test what is this compressive strength. That can actually take any value between certain range. So, this one also one example of the continuous random variable.

Now, coming to the mixed random variable. So, here say that depth of rainfall at a particular rain gauge station. So, this depth of rainfall at a particular rain gauge station, which is, gives this one, that is as a continuous random variable here. Now, if you see that, there if, there are station where most of the days, if it is the 0 rainfall for the station, then **how the** how it will take. So, for a 0, suppose that out of 365 days in a year, suppose that, there are 220 days are 0 rainfall. So, there is a probability mass that is concentrated at a particular value of depth of rainfall 0. For the non zero value, it can have some distribution, so that distribution should have shown the first part at particular value that x equal to 0, that is the depth equal to 0. Some mass is concentrated there, some

probability mass is concentrated there and rest of the probability is distributed over the positive side of the axis. So, this is an example of the mixed random variable.

(Refer Slide Time: 24:10)



The slide is titled "Special type of RV" and contains a section for "Indicator Random Variable". It lists two points: it is a special kind of random variable associated with the occurrence of an event, and an indicator random variable I_A represents outcomes in set A as 1 and outcomes outside A as 0. The slide footer includes the NPTEL logo, the text "Reliability Methods in Civil Engineering: R3.1", the name "Dr. Rajib Haity, IIT Kharagpur", and the number "11".

So, saying these three types of random variable, it is important to note that another special type of random variable, which is known as the indicator random variable. So, it is a special kind of random variable associated with the occurrence of an event. So, this random variable, which is, says that whether the, a particular event is occurred or not. Say, an indicated random variable, I_A , represents all the outcomes in a set A as 1 and all the outcome outside A as 0. So, it is like that, whether yes or no. So, if I say that again, if I take the road accident, whether there is a road accident. If there is road accident, this random variable is fixed in such a way that is, that functional correspondence that we are talking about again here.

That is, its function is that, it will just see whether the accident has taken place or not. If it is yes, it will say 1; if it is no, it will say 0. Say, the each day I am just taking this, taking this measurement and just preparing that series, so that series consist of either 1 or 0. Again for the rainy days, I will just see every day whether it is raining or not. I am not interested in this random variable, if it is designed in such a way, that it is not interested, what is depth of the rainfall. Only, the interested is that whether it is a rainy day or not. If it is not a rainy day, then it will return 0; if it is rainy day, it will return 1. So, these types of random variable just indicate the occurrence of one particular event. Then, the first case,

whether the occurrence of the event is the accident, another one is the occurrence of the rainfall or not. So, if it is yes, that is, it is within the outcome of the set, it indicates 1. If it is outside, then it indicates 0. So, this kind of random variable is known as indicator random variable.

(Refer Slide Time: 26:22)

Independence of Random Variables

- Two random variables R_1 and R_2 are independent if for all $x_1, x_2 \in \mathbb{R}$, there exist $P(R_1 = x_1 \text{ \& } R_2 \leq x_2) = P(R_1 = x_1) \cdot P(R_2 \leq x_2)$
- Random variables R_1, R_2, R_3 are *mutually independent* if for all $x_1, x_2, \dots, x_t \in \mathbb{R}$,

$$P\left(\bigcap_{i=1}^t R_i = x_i\right) = \prod_{i=1}^t P(R_i = x_i)$$

Probability Methods in Electrical Engineering: Dr. Rajib Halty, IIT Kharagpur 12

Independence of random variable, this is another important concept to understand. When we can say that two random variables are independent, then two random variables r_1 and r_2 are independent, if for all x_1, x_2 belongs to this real line. There exist that probability of r_1 equals to x_1 and the probability of r_2 less than x_2 is equals to probability r_1 equals to x_1 multiplied by probability of r_2 equals to x_2 . Then, if this condition is satisfied, we say that these two random variables are independent. Now, there are other kinds of independent also there, which is known as the mutually independent and there are k wise or the pair wise in independence. Suppose, that there are three, more than two random variables, which is the r_1, r_2, r_3 , if we will say that these are mutually independent, if the probability of this particular one, particular that is r_1 equals to x_1 and r_2 equals to x_2 and r_3 equals to x_3 and all is equals to the multiplication of the individual probabilities. Then, this can be **this can be** told that, these random variables r_1, r_2, r_3 are mutually independent. Then, there are k wise independent **independent** random variables. Now, out of this set, suppose there are n random variables are there and in that, out of that **out of that** n random variables, it will say to be k wise independent, if all the subsets of the k variables are mutually independent.

(Refer Slide Time: 28:05)

Independence of RVs...contd.

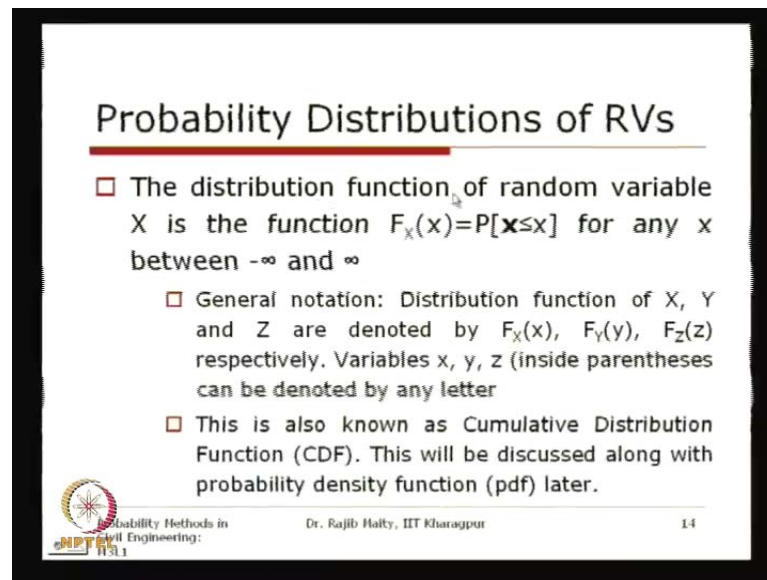
- A collection of random variables are said to *k-wise independent* if all the subsets of *k*-variables are *mutually independent*
- When the number of variables reduced to 2 and if the subsets are *mutually independent*, the random variables are said to be *pair-wise independent*

Probability Methods in Civil Engineering: Dr. Rajib Saha, IIT Kharagpur 53

So, obviously the k is less than n . So, whatever the total number of random variables are there, if I now pick up randomly k random variables and if it satisfies that the mutually independence among this k random variables, then we said those n random variables are k wise independent. Now, if this k becomes 2, then when this number of the variable reveals 2 out of n , I take the 2, that if the subsets are mutually independent. Then, these random variables are said to be pair wise independent. So, what is happening? Out of n random variables, I am picking up n random; I am picking up only two random variables.

One small correction here is that, what that r^2 is less than equals to x^2 was written; so it should be r^2 equals to x^2 . Now, the **the** most important thing, why we just learn this random variable is that probability distribution of random variables.

(Refer Slide Time: 29:20)



Probability Distributions of RVs

- The distribution function of random variable X is the function $F_X(x) = P[X \leq x]$ for any x between $-\infty$ and ∞
 - General notation: Distribution function of X , Y and Z are denoted by $F_X(x)$, $F_Y(y)$, $F_Z(z)$ respectively. Variables x , y , z (inside parentheses can be denoted by any letter
 - This is also known as Cumulative Distribution Function (CDF). This will be discussed along with probability density function (pdf) later.

Probability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 14

The distribution function of a random variable x is a function, which is denoted by this capital letter F subscript. The capital letter of that random variable, if it is x , then this is a capital letter of this x . With some smaller case letter, it may or may not be x , which can **which can** take any letter. So, which is nothing but this, inside this one, it is nothing but, as in few previous slides we have seen that this is nothing but, the specific value of that random variable. So, this is denoted as the, this is the distribution function of the random variable x .

Which is nothing but, is the probability of that random variable x , when it is less than or equal to that specific value of that random variable. So, this is the way, we define the distribution function of the random variable x , which is **which is** valid over the region from the minus infinity to the plus infinity, so, the entire real axis. So, here again, the general notation says that the distribution function of x , y and z are generally denoted by F subscript capital x any letter lowercase letter f x x , f y y , f z z respectively. This variable, that is a lower case letter, that what I was just telling here, is x inside y inside z inside the parentheses can be denoted by any letter, as these are nothing but, the specific value of that random variable. This is also known as the cumulative distribution function, when in the, most probably in the next class, we will just discuss about this cumulative distribution function. This is actually the cumulative distribution function of the, this one of this particular random variable; this will be discussed along with probability density function pdf later. So, what we want to tell is that, for a particular

random variable, this random variable over this range, some probabilities are assigned for the specific range and that **that the** how it is **how it is** distributed over this real axis is known as the distribution function of that particular random variable.

(Refer Slide Time: 31:46)

Probability Distributions of different types of RVs

- **Discrete Probability Distribution**
 - It is a mathematical function (denoted as $p(x)$) that satisfies the following properties:
 - The probability of any event x can take a specific value $p(x)$ i.e. $[P(X=x)] = p(x) = p_x$
 - $p(x)$ is non-negative for all real x
 - The sum of $p(x)$ over all possible values of x is 1
 - Though mathematically there is no restriction, in practice, discrete probability distribution function only defines for integers values

NPTEL
Probability Methods in Civil Engineering

Dr. Rajib Mallik, IIT Kharagpur

15

Now, probability distribution of different types of random variables. Now, we have discussed about three different types of random variables; the first is the discrete random variable. We will see how it is defined that it is the mathematical function denoted by generally, for when it is discrete we denote by this p_x ; this is also known as the probability mass function. We will discuss this later. So, this is denoted by p_x that satisfy the following properties. The first one is that **that the** probability of any particular event x , because this is, as it is discrete, it can any take specific value. **all any** Means, that those specific value, which is there in the feasible sample space, so that particular specific value is denoted by p_x which is nothing but, the probability of the random variable taking the specific value x , denoted by either p inside that specific value x or p subscript that value, the specific value. Obviously, from the axioms of this probability, **from the axioms of this probability** this p_x is a nonnegative, for all the real x , it can either be 0 or greater than 0.

And, the summation of this all this p_x over this possible values of x is 1. This, again from the axioms of this probability. Though mathematically, there is no restriction, in practice, discrete probability distribution function only defines for integer values. So, this is just, when we have also seen in previous slide, that some example of this discrete probability

distribution that this is generally take the integer values. But, it is not specific. It is not, mathematically there is no restriction, and it can take any **any** specific value, **the way I am way** I am defining that random variable. Second type of this random variable is the continuous random **continuous random** variable. So, this continuous distribution function for that thing, it is a mathematical, it is again a mathematical function, which is denoted by this capital F subscript, that random variable or any lower case letter, particularly this random variable lower case letter that satisfy the following property.

(Refer Slide Time: 35:28)

Probability Distributions of different types of RVs ...contd

- **Mixed Probability Distribution**
 - It is a mathematical function (denoted as $F_X(x)$) that satisfies the following properties:
 - For all x

$$F_X(x) \geq 0$$
 - It is monotonically increasing function with sudden jumps/steps
 - It is 1 at $x = \infty$ and 0 at $x = -\infty$, i.e.,

$$F_X(\infty) = 1 \text{ and } F_X(-\infty) = 0$$

Probability Methods in Electrical Engineering | Dr. Rajib Mallik, IIT Kharagpur | 27

The first is that, for any specific value of this **(0)**, should be greater than equal to 0. It is monotonically increasing and continuous function. So, here it is monotonically increasing, that means whenever it starts from and it will go and it will go on increasing and you know that this can go to maximum. So, generally for it is defined from 0 to 1. So this, it can go up, it starts from 0 and go up to 1 and it increases monotonically and it is then continuous function, as this random variable itself is continuous. So, it is 1 at x is equals to infinity and 0 at x is equals to minus infinity, that is this probability at this x is 1 and probability x minus infinity is 0.

Then, the last thing is this mixed probability distribution, which is, which can take it, for some range it can take that discrete values and some range can take continuous value. This also denoted by f_X and keeping the all other properties, all other conditions, that is the properties are same, which is obviously greater than equal to 0. Though, with the only

difference with this continuous is that, it is monotonically increasing function with sudden jumps or the steps and again the third one is again same.

Now this one, this lies the difference between this continuous probability distribution and this mixed probability distribution. Now, if you just see it here, as we are just telling that there are at some point, where the probability is concentrated, now, if I want to just see that how this cumulated accumulated over time and so, it starts from here and it will go and go up to maybe the way it is drawn, it should be assumed ∞ . So, this is the jump that we are talking about. So, this is the jump where the probability masses are concentrated.

Here it is only once, so it can be concentrated in some other range. Then, there also will be one jump. So, wherever, the probability masses concentrated for some specific value, so they are generally in this distribution function. We see that type of jump here, which is the mixed probability distribution.

(Refer Slide Time: 36:27)

Usage in Statistics

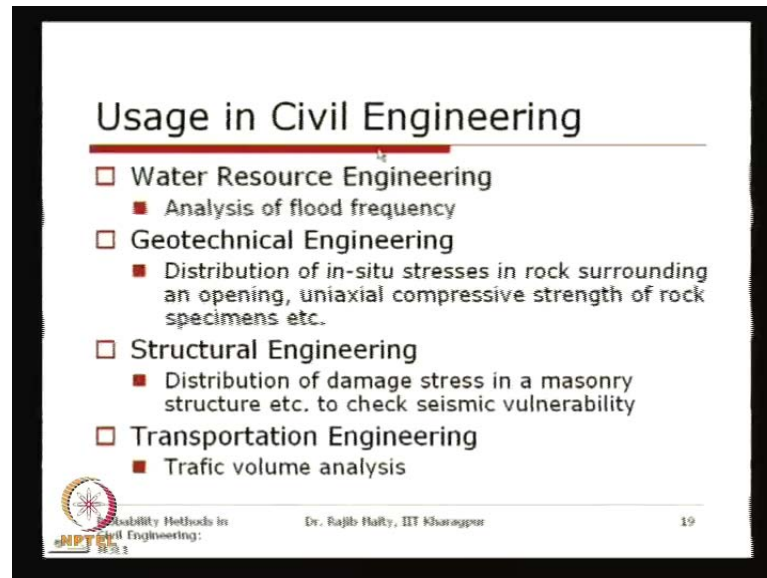
- To calculate intervals for parameters and to calculate critical regions
- To determine reasonable distributional model for univariate data
- To verify the distributional assumptions
- To study the simulation of random numbers generated from a specific probability distribution

NPTEL Probability Methods in Civil Engineering Dr. Rajib Haity, IIT Kharagpur 18

Now this, the concept, the usage of this concept of this random variable in statistics, in it is a never ending list, I should say. So, here are just few examples are given that to calculate the intervals of the parameters and to calculate critical region. We will see this, what is critical region and this interval of parameters in the subsequent classes to determine the reasonable distributional model for invariant data. Now, this data can be of any field of this civil engineering.

Now, for this analysis, the distributional model for those kinds of data, this is useful to verify the distributional assumptions. We generally, for any probabilistic model, we have assumed some distribution. Now, we have to verify whether that particular distribution is followed or not to study the simulation of random numbers generated from a specific probability distribution.

(Refer Slide Time: 37:25)



Usage in Civil Engineering

- Water Resource Engineering
 - Analysis of flood frequency
- Geotechnical Engineering
 - Distribution of in-situ stresses in rock surrounding an opening, uniaxial compressive strength of rock specimens etc.
- Structural Engineering
 - Distribution of damage stress in a masonry structure etc. to check seismic vulnerability
- Transportation Engineering
 - Traffic volume analysis

Reliability Methods in Civil Engineering: Dr. Rajib Haque, IIT Kharagpur 19

Coming to the specific civil engineering, there are different facets of civil engineering, starting from water resource engineering, geotechnical engineering, structural engineering, transportation engineering, environmental engineering and there are many such. So, in water resource, for one example, the analysis of flood frequency, this concept is used. In geotechnical engineering, distribution of in situ stresses in the rock surrounding and opening or the uniaxial compressive strength of the rock specimen, etcetera. In structural engineering, distribution of the damage stress in a masonry structure, etcetera, to check the seismic vulnerability of the structure. Transportation engineering, for example, the traffic volume analysis and this kind of thing in different application of this civil engineering.

(Refer Slide Time: 38:25)

Percentiles

- The u -percentile of a random variable X is the smallest number x_u so that $u = P[X \leq x_u] = F(x_u)$
- x_u is the inverse of the distribution function $F_X(x)$, i.e. $x_u = F_X^{-1}(u)$ within the domain $(0 \leq u \leq 1)$ and the range of x being $-\infty \leq x \leq \infty$

Probability Methods in Mechanical Engineering
Dr. Sajib Hossain, IIT Kharagpur

Now, we will see the **the** concept of the percentile for a random variable. The u percentile of a random variable x is the smallest number x_u , so that u equals to the probability of x less than u , which is equal to probability of x_u . So, to determine what is the value of this x_u , the x_u is generally the inverse of the distribution function f_x , that is, x_u is equal to the f_x inverse u , within the domain 0 to u to 1 .

So now, to know this thing, basically, if graphically if I just want to see it here, now to show this is the feasible range of that particular variable, suppose this goes on here. So I, now to calculate this percentile, what some percentile if we just say, some u , so here we can say, where is the u . Now, basically, suppose this is here, so basically, we just go and see how much percentage is covered here. So, this particular value, is graphically representing this particular value, is your u percentile of that random variable. So, generally what happens, we generally see this particular value, calculate its cumulative probability and get this one. So, that is what is your mapping, as this f_x of any value x , that is your mapping. Now, when you are coming from this side, then what we are doing, we are just giving this f_x of u inverse will give you some value that x_u , which is your x_u here.

So, this is that **inverse** function of that one to get that **get that** percentile. So, so the u percentile of a random variable x is the smallest number x_u so that u equals to probability of x less than equals to x_u . So, this is obtained that x_u is equal to inverse of that of the distribution, of that particular number, particular random variable for that

percentile u . And obviously, the u have the range from 0 to 1 and so it is expressed in percentages taken from 0 to 1 range, and the range of x being whatever **whatever** that range of particular random variable.

(Refer Slide Time: 40:55)

Properties of Distribution Functions

- Notations:
 $F(x^+) = \lim_{\epsilon \rightarrow 0} F(x + \epsilon)$, $F(x^-) = \lim_{\epsilon \rightarrow 0} F(x - \epsilon)$
 where $0 < \epsilon \rightarrow 0$
- Property 1: $F(+\infty) = 1$, $F(-\infty) = 0$
 - Proof: $F(+\infty) = P[\mathbf{x} \leq +\infty] = P(S) = 1$
 and $F(-\infty) = P[\mathbf{x} = -\infty] = 0$

Probability Methods in
 Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 21

There are different properties of this distribution functions. So we will see one by another, one by one of these properties.

In this discussion, we will follow some notation this is taken mostly from the Papoulis book. So, this is that $f(x)$ plus of this random variable x , of course, is equals to limit f that particular value x plus epsilon and $f(x^-)$ means the limit of this one, when this epsilon are greater than 0 but, it is tending to 0. So very small number, so this $f(x^+)$ means, just right side of that x and $f(x^-)$ is the just left side of that **of that** x . So, property one, the first property that is, if I take this distribution function for this infinity is equals to 1, so that the right extreme of this real x and left extreme of the real x , it starts from 0. So, it always starts from 0 ends at 1. So, to put this 1, that is $f(+\infty)$ is equals to $f(x \leq +\infty)$. So, if $f(x)$ is less than equals to plus infinity, that means it is encompassing the full sample space. So this is nothing but, that probability of the full sample spaces and we know that full sample space from the axioms of the probability, that this is equals to 1. Similarly, so if minus infinity is nothing but, the probability of x equals to $x \leq -\infty$, basically, less

than equals to minus infinity, which is a, which is basically, basically null set, so probability of the null set is equals to 0.

(Refer Slide Time: 42:40)

Properties...contd.

- Property 2: $F(x)$ is a non-decreasing function of x : if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$
 - Proof: since $\mathbf{x}(\zeta) \leq x_1$ and $\mathbf{x}(\zeta) \leq x_2$ for some ζ , $[\mathbf{x} \leq x_1]$ is a subset of the event $[\mathbf{x} \leq x_2]$. Hence $P[\mathbf{x} \leq x_1] \leq P[\mathbf{x} \leq x_2]$.
 - $F(x)$ increases from 0 to 1 as x increases from $-\infty$ to $+\infty$

Probability Methods in Civil Engineering: Dr. Rajib Maity, IIT Kharagpur 22

Second property, f is a non-decreasing function of x . That is, this non-decreasing function or monotonically increasing function, which is the word that was used in few slides previous. So, it says that, if x_1 is less than x_2 , then always this $f(x_1)$, the value of this distribution function at x_1 is less than equals to $f(x_2)$. So, it can never decrease. So, it will always, it will either be same or it will increase. So, that is basically what is known as this monotonically increasing. Proof of this one is that, if x_i is less than equal to x_1 and x_i less than equal to x_2 , for some outcome of this x_i , then x random variable less than equals to x_1 is a subset of the event x less than equals to x_2 . So what, so this x_1 is always there within this x less than equal to x_2 . If this is greater than, if this x_1 is less than x_2 , so that is why if it is a subset of this one, then obvious the probability this should be less than equals to probability of this x less than x_2 .

$F(x)$ increases from 0 to 1 as x increases from minus infinity to plus infinity. This will be 0 to 1 as x increases from minus infinity to plus infinity.

(Refer Slide Time: 44:35)

Properties...contd.

- Property 3: If $F(x_0)=0$, then $F(x)=0$ for any $x \leq x_0$
 - Proof: Since $F(-\infty)=0$, and suppose that $x(\zeta) \geq 0$ for every ζ . $F(0)=P[\mathbf{x} \leq 0]=0$ as $[\mathbf{x} \leq 0]$ is an impossible event. So, $F(x)=0$ for each $x \leq 0$

Reliability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 23

Third property says that, if x is not equal to 0, then $F(x)$ equals to 0 for any x which is less than this x_0 . So, for a specific value, if the $F(x)$ equals to 0, anything which is lower than this x_0 obviously will be 0. This is basically the same concept of that it is a non-decreasing function. So, if it is some portion, if it is 0, left side of that in the real axis context

(Refer Slide Time: 45:35)

Properties...contd.

- Property 4: $P[\mathbf{x} > x] = 1 - F(x)$
 - Proof: The events $[\mathbf{x} \leq x]$ and $[\mathbf{x} > x]$ are mutually exclusive and $[\mathbf{x} \leq x] \cup [\mathbf{x} > x] = S$. So, $P[\mathbf{x} \leq x] + P[\mathbf{x} > x] = P(S) = 1$. Then $F(x) + P[\mathbf{x} > x] = 1$, $P[\mathbf{x} > x] = 1 - F(x)$

Reliability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 24

X greater than x are mutually exclusive. So, if these two events mutually exclusive and collectively exhaustive, so collectively exhaustive means this x less than equals to x union

x greater than x is equal to full sample space s . So that, probability of this less than x and greater than x , is nothing but, probability of the total sample space. That is, the s equals to equal to 1 and now this is denoted by $f(x)$, which is this one equals to 1. So probability of x greater than x results to $1 - f(x)$.

(Refer Slide Time: 46:39)

Properties...contd.

- Property 5: The function $F(x)$ is continuous from the right: $F(x^+) = F(x)$
 - Proof: since $P[x \leq x + \epsilon] = F(x + \epsilon)$ and $F(x + \epsilon) \rightarrow F(x^+)$ when $[x \leq x + \epsilon] \rightarrow [x \leq x]$ as $\epsilon \rightarrow 0$

Probability Methods in Electrical Engineering: 19.03.1
Dr. Rajib Halty, IIT Kharagpur
25

The fifth property says, the function $f(x)$ is continuous from right. So, $f(x^+)$, that is the from right is equal to $f(x)$. Proof, since probability of x less than equals to x plus ϵ where this ϵ is nothing but ϵ is standing to 0 is equal to $f(x + \epsilon)$ and $f(x + \epsilon)$ is tending to $f(x^+)$, when this $f(x + \epsilon)$ is less than equals to $x + \epsilon$ is tending to $f(x)$ less than equal to x , as this ϵ is tending to 0. So, that is why, this is from this right hand side, if we say this is continuous from the right hand side of any specific value x .

(Refer Slide Time: 47:31)

Properties...contd.

□ Property 6: $P[x_1 < X \leq x_2] = F(x_2) - F(x_1)$

■ Proof: $[X \leq x_1]$ and $[x_1 < X \leq x_2]$ are mutually exclusive and again $[X \leq x_2] = [X \leq x_1] \cup [x_1 < X \leq x_2]$.
So,
 $P[X \leq x_2] = P[X \leq x_1] + P[x_1 < X \leq x_2]$ or

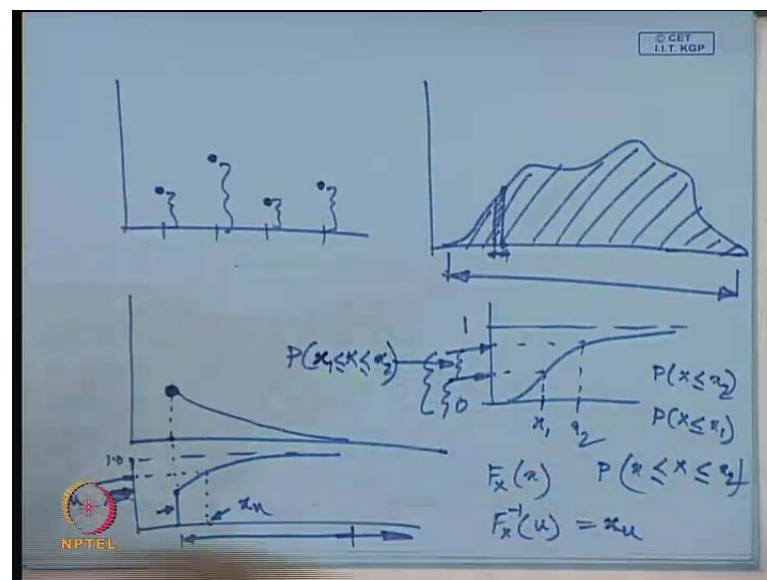
$$P[x_1 < X \leq x_2] = P[X \leq x_2] - P[X \leq x_1] = F(x_2) - F(x_1)$$

© CEE I.I.T. KGP

Probability Methods in Mechanical Engineering: Dr. Rajib Maiti, IIT Kharagpur 26

The sixth property says, for a random variable, if it is bounded by this x_1 and x_2 , the probability of the value from starting from x_1 to x_2 is equal to the probability of x_2 minus x_1 .

(Refer Slide Time: 47:53)



So, graphically if I just see it again here, that is, if I just take two values, if this is your say starts from 0 and goes like this. So, if I want to know that, what if this is your x_1 , this is x_2 , then, these probabilities that is, x_2 , so probability that x less than equals to x_2 is nothing but, this particular value. Now, probability of x less than equals to x_1 is nothing

but, this particular value. Now, if I want to know that, if this x is in between these two, x_2 and x_1 , then this probability is nothing but, whatever total probability this minus this probability. So, this is the probability that we are talking about, which is nothing but, the probability of x_1 less than equals to x less than equals to x_2 . So, this is how we get the probability for **for** a range.

Which are **which are** again, this proof **is proof** says that **that** x less than equals to x_1 and this x_1 less than x less than x_2 are mutually exclusive. Again, x less than equals to x_2 is equals to x less than equals to x_1 union x_1 to x_2 . **these** So, the probability x less than equals to x_2 is probability x less than equals to x_1 plus probability x in between x_1 to x_2 . Or, if I just take this one here, then probability of x_1 , this x random variable between x_1 to x_2 is equals to probability of x less than x_2 minus probability of $x < x_1$, which is again nothing but, this $f(x_1)$ $f(x_2)$.

(Refer Slide Time: 49:51)

Properties...contd.

□ Property 7: $P[x=x] = F(x) - F(x^-)$

■ Proof: Putting $x_1 = x - \epsilon$ and $x_2 = x$ in Property 6, we get:

$$P[x - \epsilon < x \leq x] = F(x) - F(x - \epsilon). \text{ Now taking } \epsilon \rightarrow 0$$

$$P[x=x] = F(x) - F(x^-)$$

Probability Methods in Civil Engineering: Dr. Rajib Halty, IIT Kharagpur 27

Now, the property seven says, that if probability of x is equals to a particular value, x is equals to probability of x minus probability of x just left to that one. So, at a particular point, the probability there, now these **this** properties are in general for the discrete continuous. So, for a particular point, the probability says, that at that particular point and just left to that, whatever the probability is there, so it will be like this; from that particular side to just to the left of this one. Proof, putting that x_1 equals to x minus x_i and x_2 equal to x in this property six, then we can say that probability of x minus x_i less than

x less than that x equals to probability x equals to probability x minus x_i . Now, taking this x_i tending to 0, then we can say that probability at a particular specific value is that point, probability at that point minus immediate previous value to that one, that x_i particular value.

So, if it is a discrete random variable, then just x_i left to this value, this x_i value comes to 0. So, that for a discrete value at particular x_i point, the probability is equal to that x_i the functional value at that particular point.

(Refer Slide Time: 51:25)

Concluding Remarks

- Random variable is not a variable rather a function which map all the feasible outcome of an experiment on the real line or a set of real numbers
- A complex random variable can only be define by joint distribution of real and imaginary variables
- Random variables can be either discrete or continuous if the set of events defined by the variable is either finite or infinite numbers respectively. Mixed distributions are the combination of both discrete and continuous distribution.
- Distribution of random variable with specific application to different civil engineering related problems will be discussed in the next lecture.

Dr. Rajib Maity, IIT Kharagpur

Finally, in this lecture, we have seen that random variable is not a variable rather a function which map all the feasible outcome of an experiment on the real line or a set of real numbers. A complex random variable can only be defined by the joint distribution of the real and the imaginary variable, that is, if that x plus y will be equal to that, the joint distribution of this x and x and y , both the random variables. Random variables can be either discrete or continuous if the set of events defined by this variable is either finite or infinite numbers respectively. Mixed distributions are the combination of both discrete and continuous distribution. Distribution of random variable with specific application to this different civil engineering related problems will be discussed in the next lecture. I will meet that concept of this probability density function and cumulative distribution function. Thank you.