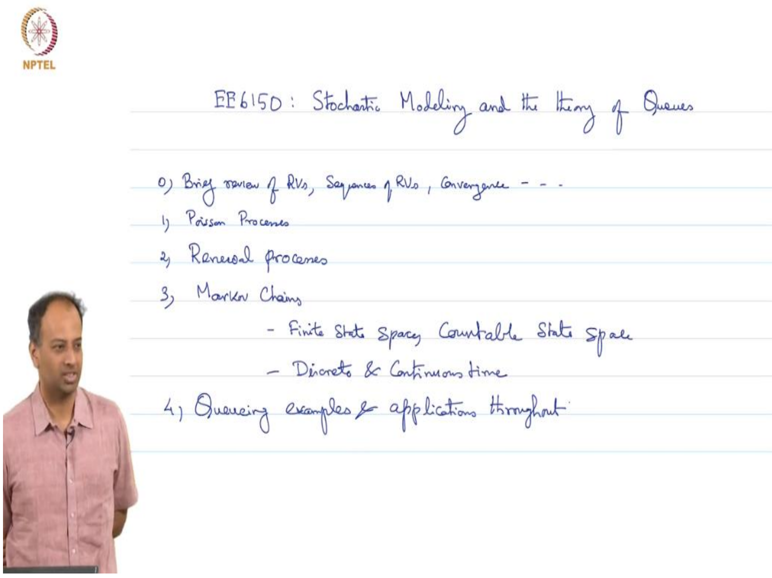
Stochastic Modeling and the Theory of Queues
Prof. Krishna Jagannathan
Department of Electrical Engineering
Indian Institute of Technology - Madras

Module - 1
Lecture - 1
Review of Probability Theory: Random Variable

(Refer Slide Time: 00:15)



Welcome. This is EE6150, Stochastic Modeling and Theory of Queues. This is a course on Discrete Stochastic Processes. In particular, the topics we cover will be as follows.

0. Brief review of Random variables, Sequences of Random variables, and Convergence.

   Then the main topics in this course are the following:

1. Poisson Process.
2. Renewal processes.
3. Markov chains {In detail, about half the course is about Markov chains}. In Markov chains, we will learn
   a. Finite State Space and Countable State Space. In both,
   b. Discrete and Continuous Time.

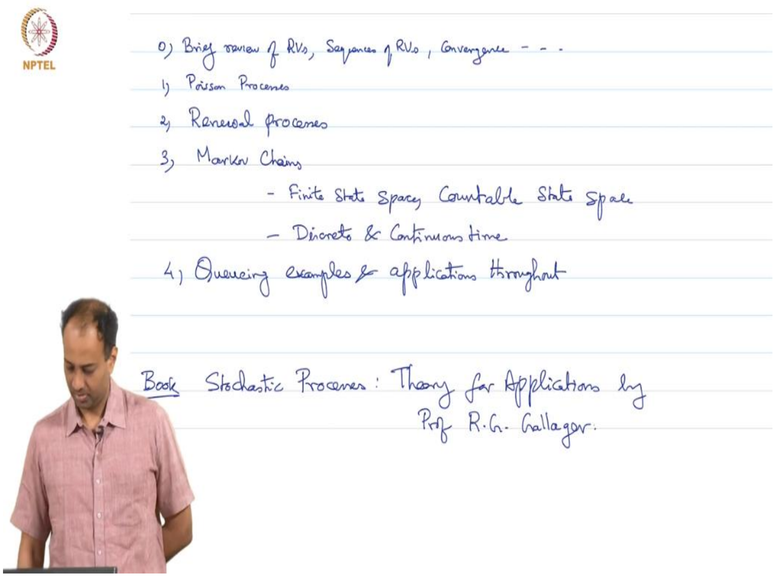
   And we will look at

4. Queueing Examples and Applications throughout.

The word 'queue' is the only word in English language that is still pronounced the same way when the last four letters are removed. *Queueing*: is British, *Queuing* is American. The word **Queueing** is one of the few words in English which has 5 vowels consecutively. Learn more

This is essentially a discrete stochastic processes course with particular emphasis on queueing applications and examples. We will also give examples from other fields occasionally.

(Refer Slide Time: 02:53)



I will closely follow a book, **Stochastic Processes, Theory for Applications** authored by Gallager. I learned from this book, so I find it easy to teach from this book. It is a good book especially for applied people. So, if you are not a pure mathematician and if you want to look at engineering applications, this is a good book to learn from. It is sufficiently rigorous without being too fussy about details. And it is good for application, certainly.

(Refer Slide Time: 04:05)

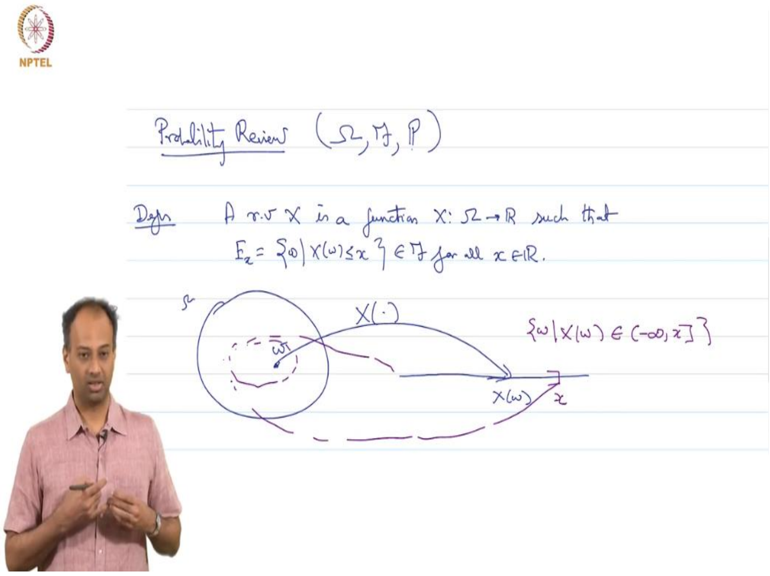


Figure 1

Now, let’s start with basic probability review.

All of probability theory starts with a triple $(\Omega, \mathcal{F}, \mathbb{P})$. This you would have seen in a probability course. This Ω here, is a sample space consisting of all possible outcomes of a random experiment. This $\mathcal{F}$ is a collection of events. It consists of subsets of Ω, those are of interest. This $\mathcal{F}$ is known as a **sigma algebra**, sigma algebra of events. And this $\mathbb{P}$, is a probability measure which assigns numbers between 0 and 1 to the events.

So, each element of $\mathcal{F}$ is some subset of Ω, and probability is assigned to each of these events as a number between 0 and 1. These probabilities have to satisfy some axioms, in particular

1. The probability of the empty set is always 0.
2. The probability of sample space is always 1.
3. The probability of a countable union of disjoint events is equal to the sum of the probabilities of the individual.

So, these 3 axioms have to be satisfied. These may be familiar from a probability course.

Now, everything we will do, assumes some implicit probability space.

There is some Ω, $\mathcal{F}$, and $\mathbb{P}$ which is already been defined and given to us. And everything starts from there. Once you have this probability space, there are random variables which can be defined on Ω.

Q. What is a random variable?

A: Random variable is a function. Random variable is a function from $\Omega \rightarrow \mathbb{R}$.

So, a random variable is a terrible **misnomer**, it is neither **random**, nor it is a **variable**. It is a deterministic function from $\Omega \rightarrow \mathbb{R}$.

Q: Then, what is random?

A: The input to the random variable. When the random experiment occurs, there is a ω that is picked by nature, which is the random part.

That is the only random thing in all of probability theory. Once ω realizes $X(\omega)$ is a fixed real number, it is ω that is random, it is not the random variable that is random.

Think of it this way, let us consider a function $f(x) = x^2$. It is not random.

But if x is random, $f(x)$ will be random. So, the randomness cannot be blamed on the function.

It is to be blamed on the input to the function. So, a random variable is a function from $\Omega \rightarrow \mathbb{R}$. There is a mild technical condition that is required; it is not just any function. Random variable is a function from $\Omega \rightarrow \mathbb{R}$ such that, let us say, if you look at those ω's which map to less than or equal to x ; These ω's have to be events, events for all real x.

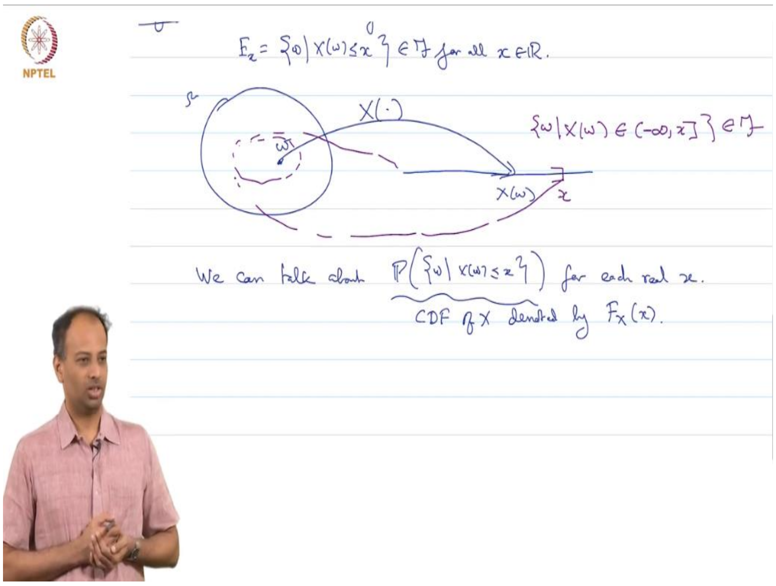
This is a definition of a random variable. We can understand this better by drawing this picture [Figure 1]. The ω is point inside the circle, the circle is the set of all possible values of ω which is the sample space Ω. And there is some $\mathcal{F}$, $\mathbb{P}$ defined on this Ω. So, every time; so, there is a ω that is picked by nature. X maps ω on to the real line, and thus for every ω, $X(\omega)$ realises as a real number.

Now for some x fixed on the real line and for an interval from $-\infty$ to x. We are looking at the set of all ω's for which $X(\omega)$ belongs to $(-\infty, x]$. In some sense you can look at this as, like the pre-image of the semi-infinite interval, $(-\infty, x]$, under this mapping. You look at those ω's which map to less than equal to x. There are some ω's which map to less than equal to x, there are other ω's which map to greater than x.

You look at those ω's which map to less than or equal to x. So, that will be some subset of the sample space obviously. You want that to be an event for every real x.

A function which satisfies this property is called a **measurable function**. In probability, it is just called a **random variable**.

(Refer Slide Time: 10:11)



Reason: Why do we want to impose this condition?

Suppose the condition is satisfied, for every x, $X(\omega) \leq x$ is an event.

$\{\omega | X(\omega) \in (-\infty, x]\} \in \mathcal{F}$. Then you can talk about the probability of the event. Once some subset of Ω is an event, you can assign probability to it. So, you can talk about this for each real x.

If this were not an event, then you cannot talk about its probability. Only events have probabilities. Not all subsets of ω may have probabilities assigned to them. Now, $\mathbb{P}(\{\omega|X(\omega) \leq x\})$ is very important in probability. What is this? This is the CDF. This is the cumulative distribution function of X. It is denoted by $F_X(x)$. So, $F_X(x)$ is simply the probability of random variable taking values less than or equal to x.

It is easy to establish that, if you are given the CDF of a random variable, then you know everything about its statistical properties. So, for any nice subset of $\mathbb{R}$; you can talk about the random variable taking values in that subset of $\mathbb{R}$. By nice, I mean **Borel**, Borel subset of $\mathbb{R}$. Now, this was a discussion about one random variable. We could of course have multiple random variables on the same sample space $(\Omega, \mathcal{F}, \mathbb{P})$.

(Refer Slide Time: 12:20)

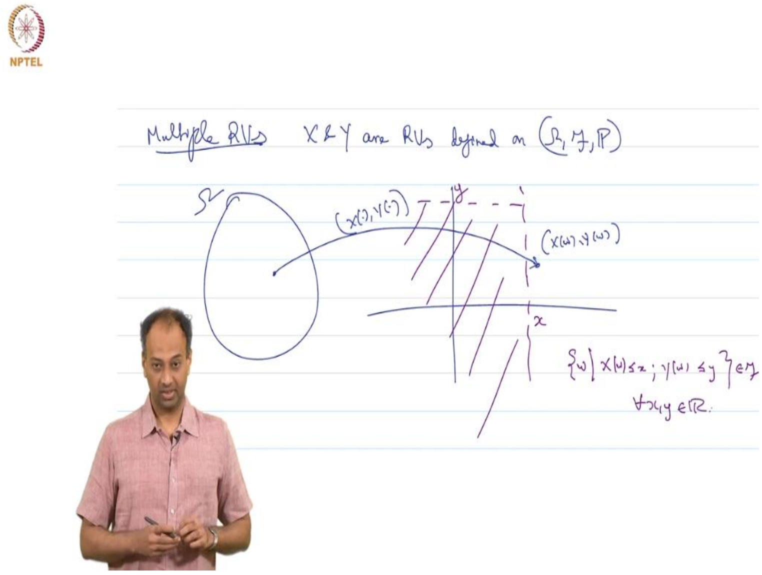


Figure 2

If you want to discuss that, Multiple Random Variables. Let us talk about two random variables first. Let us say X and Y are random variables are defined on $(\Omega, \mathcal{F}, \mathbb{P})$. There are two different random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. i.e., X is a function from $\Omega \to \mathbb{R}$, satisfying the property that we talked about; and Y is another function from $\Omega \to \mathbb{R}$, satisfying the same property.
So, the pre-images of these semi-infinite intervals are events, for both X and Y. The values that these random variables take is determined by the same underlying outcome, the elementary outcome, ω. So, you can think of X and Y as, let us say, one of them is rainfall, the other is humidity on a particular day.

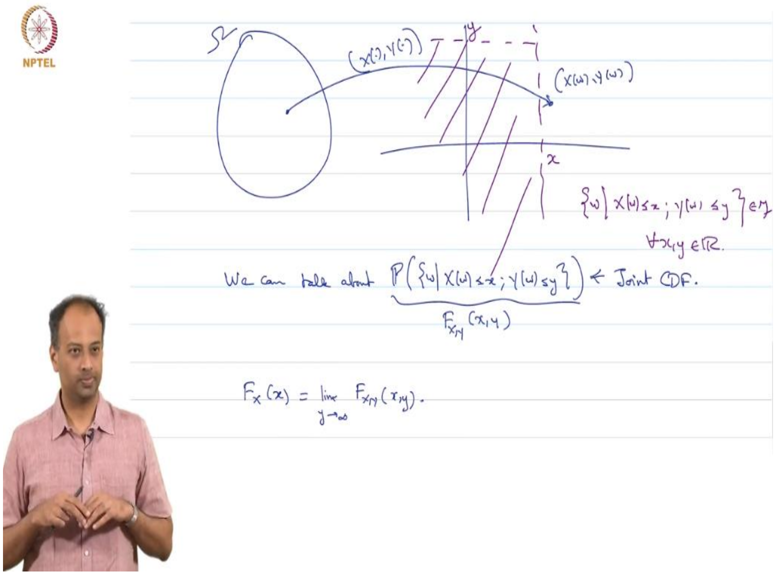
There is an underlying randomness which gives both, the values for both X and Y. So, the way to look at this in Figure 2. Observe it using two axes and look at $\mathbb{R}^2$ for the corresponding point inside the circle. So, X is a map from $\Omega \to \mathbb{R}$; Y is a map from $\Omega \to \mathbb{R}$. It is useful here to look at the joint map X, Y from $\Omega \to \mathbb{R}^2$, because it is the same ω which realises the values of X and Y. So, you look at the map from the sample space to $\mathbb{R}^2$.

And then, you realise a point $\big(X(\omega), Y(\omega)\big)$. For every ω, there is a point that realises on $\mathbb{R}^2$. This reason we look at it this way as opposed to looking at two real lines, (one for X and one for Y) is that it is the same underlying ω which realises these values. So, there could be some sort of correlation or dependence between the values it takes. So, a higher temperature could imply a higher humidity or whatever it is, that the underlying model is talking about.

So, instead of looking at two real lines, we look at $\mathbb{R}^2$; which captures the joint variation in these two random variables. Now, what can be shown is that, just like you have a CDF in the one-dimensional case, you can argue that if you have the semi-infinite rectangles, as shown in Figure 2. We will look at the shaded region for some x and y on the axes respectively.

Consider the ω's for which $X(\omega)$ is less than or equal to x and $Y(\omega)$ less than or equal to y. We can argue that this is an event. And if this is an event for all x, $y \in \mathbb{R}$;, you can talk about its probability. And what is the probability of that?

(Refer Slide Time: 16:13)



So, we can talk about the probability, that $X(\omega)$ is less than or equal to x and $Y(\omega)$ less than or equal to y; $\mathbb{P}(\{\omega|X(\omega) \leq x; Y(\omega) \leq y\})$, which is the joint CDF, denoted by $F_{X,Y}(x, y)$. **Convention to followed**: Always big letters for the names of the random variables and small letters for the specific real values that they take. Please do not interchange this.

So, again, if you are given the joint CDF of X and Y, you know everything about these joint statistical properties.

In fact, given the joint CDF, you can calculate the marginal CDF. So, for example, if you put; in joint CDF, they have several properties, but if you want to calculate $F_X(x)$, just the marginal CDF of X, you can just calculate it by

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y)$$

Then you will get the marginal of X. And similarly, if you want the marginal of Y, send the other variable x to infinity.

Given the joint CDF, you can always recover the marginal CDF. But of course, given the marginal CDF, there is no way you can recover the joint CDF, because you are losing information on how they vary jointly. You can only understand, how x takes values on real line and y takes values on its real line, on the other real line. You miss the joint variation. So, the information is lost.

But once you know the joint CDF, you know everything about the statistical properties of X and Y. Similarly, this can be extended to n random variables. Nothing changes; it is just a map from $\Omega \rightarrow \mathbb{R}^{n}$, where $\mathbb{R}^{n}$ is the n-dimensional Euclidean space. This is the end of the first module.