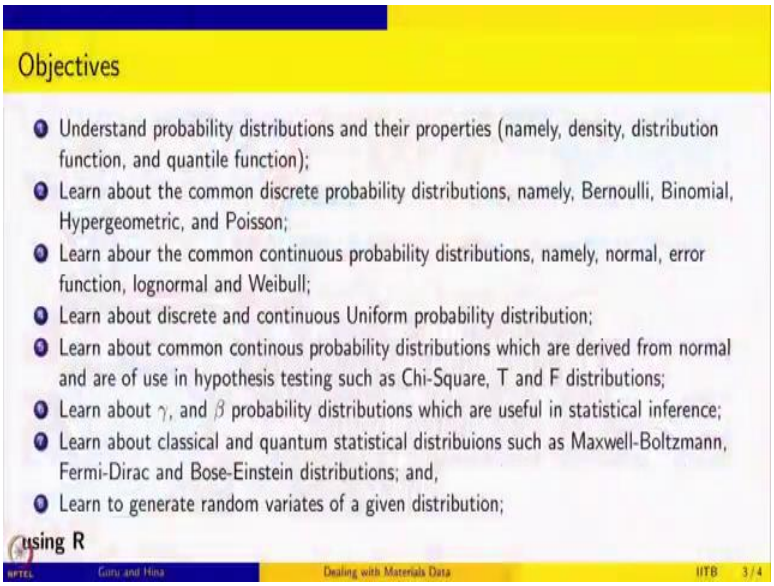
Dealing with Materials Data: Collection, Analysis and Interpretation
Professor M P Gururanjan
Professor Hina A Gokhale
Department of Metallurgical Engineering and Materials Science
Indian Institute of Technology, Bombay
Lecture 37
Probability distributions.

Welcome to Dealing with Materials Data. This course is on collection, analysis and interpretation of materials data. We have completed two modules, first one was the introduction to R, the second one was descriptive statistics using R and in this module we are going to learn about probability distributions and how to use R to understand probability distributions and work with them.

(Refer Slide Time: 0:41)



So, this is the objective of this module. We want to understand probability distributions and their properties. The properties include density, distribution function and quantile function. We want to learn about common discrete probability distributions such as Bernoulli, Binomial, hypergeometric and Poisson, for example. And we want to learn about common continuous probability distributions like normal, error function, lognormal and Weibull, for example.

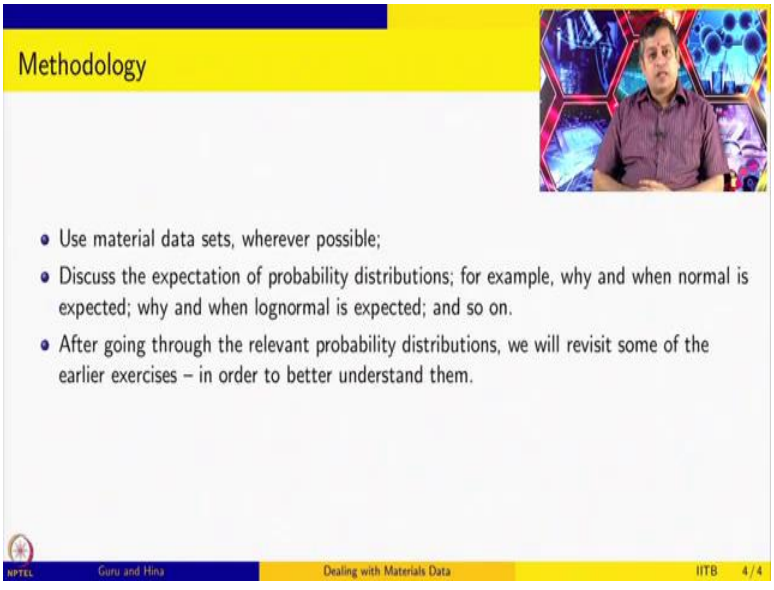
We want to learn about discrete and continuous uniform probability distribution so uniform probability distribution can be either discrete or continuous, so we want to learn about uniform probability distribution, because it plays a key role in some of the simulations that we have been

carrying out and in order to understand what this simulations are its important to understand uniform probability distribution.

We all also learn about some common continuous probability distributions which are derived from normal such as Chi square, T and F distributions. Because these distributions are very useful in making parameter estimation or in understanding hypothesis testing, so they are of great use in hypothesis testing, for example, so we will learn about them. We will also learn about gamma and beta probability distributions so they have some use and relevance in inference so we will learn about them.

And from material science we do know that there are classical and quantum statistical distributions, Maxwell Boltzmann is a classical one, Fermi Dirac and Bose Einstein are the quantum statistical distributions, so we will have some exposure to them also. In the same schema of things what are these distributions and how do they look and things like that we will learn. And we will also learn how to generate random variates from any of this distributions for doing computations and calculations. All of this we want to do using R and of course using materials data.

(Refer Slide Time: 2:45)



So, we will use materials data sets wherever possible, our examples will be from materials wherever possible and we will also discuss, what is the expectation in terms of probability distributions? When and why one should use normal? When and why one should use lognormal

and so on. So, there are cases where we know a priori that the best distribution to describe the given data is Weibull, for example, or Hypergeometric, for example.

So, we will discuss some of this cases and so we will understand these issues better and also how to deal with them using R or how to work with this using R. And one more thing we are going to do is that even though we are doing probability distribution now, we have used some of these ideas in previous module for example. So after going through this probability distribution session we are going to revisit some of the exercises to better understand what we did earlier.

For example, we have try to fit the empirical data to some of distributions so after understanding this distributions we will go back and see what the fitting exercise was and what did we get and how do we understand those results a little bit better. So, that is what we are going to do in this module so this is the third module on probability distributions using R, thank you.