**Probability Theory: Foundation for Data Science**

**Descriptive Statistics and the Axioms of Probability**

Understand the foundation of probability and its relationship to statistics and data science. We’ll learn what it means to calculate a probability, independent and dependent outcomes, and conditional events. We’ll study discrete and continuous random variables and see how this fits with data collection. We’ll end the course with Gaussian (normal) random variables and the Central Limit Theorem and understand it’s fundamental importance for all of statistics and data science.

**3 videos•Total 51 minutes**

* Intro to Probability•13 minutes
* Axioms of Probability•19 minutes
* Counting: Permutations and Combinations•17 minutes

**2 readings•Total 20 minutes**

* Intro to Probability•10 minutes
* Introducing the formula sheet for this course•10 minutes

**1 quiz•Total 45 minutes**

* Homework: Descriptive Statistics and the Axioms of Probability•45 minutes

**Conditional Probability**

The notion of “conditional probability” is a very useful concept from Probability Theory and in this module we introduce the idea of “conditioning” and Bayes’ Formula. The fundamental concept of “independent event” then naturally arises from the notion of conditioning. Conditional and independent events are fundamental concepts in understanding statistical results.

**2 videos•Total 52 minutes**

* Conditional Probability and Bayes Theorem•24 minutes
* Independent Events•27 minutes

**1 reading•Total 60 minutes**

* Conditional Probability and Bayes Theorem•60 minutes

**1 quiz•Total 30 minutes**

* Homework: Conditional Probability•30 minutes

**Discrete Random Variables**

The concept of a “random variable” (r.v.) is fundamental and often used in statistics. In this module we’ll study various named discrete random variables. We’ll learn some of their properties and why they are important. We’ll also calculate the expectation and variance for these random variables.

**4 videos•Total 77 minutes**

* Discrete Random Variables•20 minutes
* Bernoulli and Geometric Random Variables•11 minutes
* Expectation and Variance•21 minutes
* Binomial and Negative Binomial Random Variables•24 minutes

**1 reading•Total 60 minutes**

* Discrete Random Variables•60 minutes

**1 quiz•Total 45 minutes**

* Homework: Discrete Random Variables•45 minutes

**Continuous Random Variables**

In this module, we’ll extend our definition of random variables to include continuous random variables. The concepts in this unit are crucial since a substantial portion of statistics deals with the analysis of continuous random variables. We’ll begin with uniform and exponential random variables and then study Gaussian, or normal, random variables.

**4 videos•Total 96 minutes**

* Continuous Random Variables•21 minutes
* The Gaussian (normal) Random Variable Part 1•19 minutes
* The Normal Random Variable Part 2•26 minutes
* The Poisson and Exponential Random Variables•29 minutes

**2 readings•Total 120 minutes**

* Continuous random variables•60 minutes
* Normal Random Variable•60 minutes

**1 quiz•Total 30 minutes**

* Homework: Continuous Random Variables•30 minutes

**Joint Distributions and Covariance**

Module details

The power of statistics lies in being able to study the outcomes and effects of multiple random variables (i.e. sometimes referred to as “data”). Thus, in this module, we’ll learn about the concept of “joint distribution” which allows us to generalize probability theory to the multivariate case.

**3 videos•Total 66 minutes**

* Covariance and Correlation•34 minutes
* More on Expectation and Variance•15 minutes
* Jointly Distributed Random Variables•15 minutes

**1 reading•Total 60 minutes**

* Covariance and Correlation•60 minutes

**1 quiz•Total 30 minutes**

* Homework: Joint Distributions and Covariance•30 minutes

**The Central Limit Theorem**

Module details

The Central Limit Theorem (CLT) is a crucial result used in the analysis of data. In this module, we’ll introduce the CLT and it’s applications such as characterizing the distribution of the mean of a large data set. This will set the stage for the next course.

**2 videos•Total 40 minutes**

* Introduction to the Central Limit Theorem•20 minutes
* Central Limit Theorem Examples•19 minutes

**1 reading•Total 60 minutes**

* Central Limit Theorem•60 minutes

**1 quiz•Total 30 minutes**

* Homework: Central Limit Theorem•30 minutes