Guide to Lectures: Lecture 1

This lecture reviews key concepts and techniques from basic probability. It is important to note that this (and lecture 2) are not intended to replace a course in probability theory and a course in mathematical statistics. Rather, it is intended to introduce all of the probability and statistics concepts that will be used in regression and time series modeling in this course. The mathematical notations and techniques used in these two lectures covered in Wooldridge's text Appendix A.

The first two lectures of this course are mathematical oriented; the students are asked to perform paper-and-pencil calculations. On the other hand, the remaining 12 lectures are empirical oriented with a lot of empirical applications, where most of the questions in the weekly assignments and three labs require the use of *R*. Nevertheless, all of the statistical models are first introduced in mathematical formulation and all of the underlying statistical assumptions are emphasized throughout.

Pay attention to the language and terminology used in this course. We follow the traditional language used in statistics and econometrics. Some of them are different from those used in the machine learning field. For instance, the word "features" in this course is used in the English sense; it is not used to mean "explanatory variables" in a regression, time series, or machine learning models. As a data scientist, you will have to learn different terminologies used in various fields such as computer science, engineering, statistics, mathematics, etc. So, whenever you are not sure about the language or particular terminology being used, please ask.

Finally, assumptions underlying each of the probability models and statistical models are important. When you are reading the texts and watching the video lectures, pay special attention to all of the assumptions made when a model or technique is introduced. Later in the course, we will examine the consequence and potential remedies when one or more underlying statistical assumptions fails. In addition, regression modeling are the modeling of a conditional expectation function, as you read the assigned readings this week, make sure you understand all of the explanations related to and the properties of conditional expectation.

Main Topics Covered in this Lecture 1:

- Probability Space, Area Diagram
- Marginal, Joint, and Conditional Probability
- Bayes' Rule
- Random Variables
- Functions of random variables: Expectation, Variance, and Covariance
- Properties of Expectation, Variances, and Covariance
- Variance-Covariance Matrix
- Conditional Expectation
- Parameters and Estimators
- Properties of estimators: unbiasedness, consistency, efficiency, asymptotic normality
- Method of Estimation: Ordinary Least Square (OLS), Method of Moments (MOM), and Maximum Likelihood Estimation (MLE)

Learning Outcomes:

After successfully completing this lecture, you will be able to

- define basic probability concepts, such as random variables, marginal, joint, and conditional probability, and to apply them in formulating data analytical problems
- understand the features of probability distributions - expectation, conditional expectation, variance, covariance, and correlation - and their properties
- illustrate the difference between population and samples
- distinguish among parameters, estimators, and estimates
- define random sampling
- explain the three common approaches to parameter estimation

Reading 1: W2012: Appendix B: Fundamentals of Probability

When reading this chapter, ask yourself the following questions:

- What is a probability space?
- What is a random variable? What is a continuous random variable? What is a discrete random variable?
- Can you give some examples of continuous random variables and discrete random variables?
- What is relationship between marginal, conditional, and joint distributions?
- What are the key properties of expectation, variance, conditional expectation, covariance, and correlation?
- What is the population mean (variance, covariance, etc)?

Whenever you encounter expectation (variance, etc), try to express them in integral or summation form. Although we did not spend much time in the video lecture, spend time reading *B.5: The Normal and Related Distributions* because these distributions play a crucial role in regression modeling and statistical inference. Try to derive the formulas in this chapter to ensure that you really understand these concepts.

Reading 2: W2012 Appendix C1-C4: Fundamentals of Mathematical Statistics

When reading this chapter, ask yourself the following questions:

- What is a population (in statistical sense)?
- What are parameters? What is parameter estimation? What is a point estimate? What is an interval estimate?
- What is random sampling? What does it mean when a set of random variables are *independent and identically distributed*?
- Is an estimator a random variable? Why? or why not?
- What is unbiasedness? What does it mean when we say an estimator is an unbiased estimator? Does it make sense to say "an estimate is unbiased"? Why? or why not?
- What is the sampling variance of an estimators?
- What is an efficiency estimator? What is mean square error (MSE)?
- What is a consistent estimator? How is it different from an unbiased estimator? Can an estimator be both unbiased and consistent? Can an estimator be biased but consistent? Can an estimator be unbiased but inconsistent? Can you give some examples?
- What is asymptotic normality?
- What does the Law of Large Number (LLN) mean? What are the assumptions associated with LLN?

- What does Central Limit Theorem (CLT) say? What are the assumptions of this theorem?
- What is the method of moment (MOM) estimation? Ordinary Least Square (OLS) estimation? Maximum Likelihood (MLE) Estimation?