Guide to the Lecture: Lecture 2

This lecture reviews some of the most fundamental concepts in mathematical statistics, with an emphasis on the *Neyman-Pearson Approach* of hypothesis testing, the most common use "logic" to statistical inference. It also covers both the frequentist approach and (very briefly) the Bayesian approach, the notions of objective and subjective probabilities, their roles in hypothesis testing, p-value, type-I error, type-II error, the power of a test, potential corrections of p-value (in the case of multiple tests), size of a test (or as Fisher's termed it, "statistical significance"), family-wide error rates, stopping rules, Bayes' Theorem, and the likelihood function. Many examples are given in the asynchronous lecture.

This lecture is not mathematical; it is, however, rather abstract, conceptual, and philosophical. Nevertheless, the *Neyman-Pearson approach* to hypothesis testing and related concepts play a central role in modern statistics, and they will be used throughout the entire course (and possibly in any course in this program in which classical statistical techniques are used). As such, study very carefully the assigned reading. You most likely will have to re-visit this chapter multiple times in the rest of the course.

As mentioned in the *Guide to Lecture 1*, it is important to note that the first two lectures are not intended to replace a rigorous sequence of courses in probability theory and mathematical statistics. Rather, it is merely intended to introduce all of the probability and statistics concepts that will be used in regression and time series modeling in this course.

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 applications, where most of the questions in the weekly homework and the three labs require the use of *R*. Nevertheless, all of the 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 or time series model. 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. Since we will be using statistical inference extensively throughout the course, the concepts and “philosophy” covered in this lecture is extremely important. Discuss with your classmates to enhance your understanding and post questions to the Wall where needed.

Main Topics Covered in this Lecture:

- The Neyman-Pearson approach and all of the related concepts (such as Type-I error, Type-II error, the power of a test, etc.

- Multiple comparisons
- Planned vs. posthoc tests
- Stopping rules
- Bayesian Approach to Statistical Inference
- Comparison of Frequentist and the Bayesian Approach to Statistical Inference
- Prior and posterior distributions
- The likelihood principle

Learning Outcomes:

After successfully completing this lecture, you shall be able to
- understand and apply the Neyman-Pearson approach to statistical inference
- explain the subjective and objective definitions of probability
- list the axioms of probability
- define concepts such as null and alternative hypotheses, rejection region, size of the test, Type I error, Type II error, the beta, and power of a test
- describe multiple comparisons, post-hoc tests, and planned tests
- understand the meaning of statistical significance
- define objective probability and its relationship (if at all) to hypothesis
- apply Bayes' Theorem
- describe the Bayesian approach to statistical inference

Reading 1: *Understanding Psychology as a Science*, Palgrave, 2008. Ch.3
- You may want to study this chapter in several sittings, perhaps even re-read the entire chapter at least a couple of times. Despite you probably have seen and used hypothesis testing and other statistical inferences for many years, the concepts covered in this chapter may not be familiar to you. Even though it is used implicitly in many undergraduate statistics textbooks, the Neyman-Pearson approach to statistical inference is not well-explained at the undergraduate level.

When reading this chapter, ask yourself the following questions:
- What is the Neyman-Pearson logic to hypothesis testing, and how is it difference to your previous understanding of hypothesis testing?
- What does probability mean under the Neyman-Pearson paradigm?
- What is "significance" in Fisher's term? What is the "size of the test" in Neyman's term? What is the relationship of these two concepts?
- What is a rejection region? How is it related to the underlying probability distribution?
- Whenever we are conducting hypothesis testing, are we just testing a number (called an estimate) against another number (called the true parameter value), or is there more to it? (Hint: think about the role of the underlying probability distribution in the context of hypothesis testing.)
- What is type-I error, type-II error, alpha, and beta in the context of hypothesis testing?
- Can one control both type-I error and type-II error in a hypothesis testing? If so, prove it. If not, explain your rationale.
- What is the power of a test? Given two test, would you prefer the test with a higher or lower power? What is the tradeoff?

- What are the major criticism of the Neyman-Pearson approach to hypothesis testing?

In addition to the questions above, when watching the asynchronous lecture, ask yourself the following questions:

- What is the main difference between frequentist and Bayesian statistics?
- What is and what is not the p-value in the context of a hypothesis testing?
- What is the multiple comparison problem?
- What is family-wise error rate and under what situation should it be used?
- How is the planned comparison different from that of post-hoc comparison?
- What is Bayes' rule, what elements does it include, and how is it applied?