Statistics, Fall 2021 統計學暨實習

Jui-Chung Yang 楊睿中   TA:
Wed, 9:00 – 12:00, location: TBA   TBA
Office Hours: CSS 755, Wed, 12:00 – 13:00.   TBA

> *Data is the new oil. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value* (Clive Humby, 2006).
> *In school, we rarely learn probability & statistics, leaving us victims of our mind's own inability to process random events.* (Neil deGrasse Tyson, 2017).

This is the first course in statistics, *a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data* (Merriam-Webster's Dictionary,) for second-year economics majors. In this course, I will focus on basic topics such as *probability*, *point/interval estimation*, and *hypothesis testing*. Other topics, *e.g.*, *linear regression* and *statistical learning*, will be covered in the next course in spring.

Unlike most statistics courses in Taiwan, students at National Taiwan University are required to learn `R`, a programming language and free software environment for statistical computing and graphics. Some basic coding knowledge is a plus but not required. In this course, I will teach `R` from the basics.

In the age of big data, the statistical methodology has been expanding breathtakingly. I understand that students at NTU are eager to learn state-of-the-art statistical methods. However, a wise man once said *we must learn to walk before we can run*. In this semester we will focus on *walking*. If you want to know more about *running*, please feel free to come to my office hours, or see *supplemental reading*.

Required Reading

- Linton, Oliver (2017), *Probability, Statistics and Econometrics* Academic Press

- Heumann, Christian, Michael Schomaker, and Shalabh (2016), *Introduction to Statistics and Data Analysis: With Exercises, Solutions and Applications in* `R`, Springer.

Supplemental Reading

Casella and Berger (2001) is a classic for students devoted to statistics. James, Witten, Hastie, and Tibshirani (2013) is an excellent pratical treatment of statistical learning algorithms. Both Kuan (2004) and Chen (2019) are excellent textbooks in Chinese.

1. Casella, George and Roger L. Berger (2001), *Statistical Inference*, 2/E, Cengage Learning.

2. James, G., D. Witten, T. Hastie, and R. Tibshirani, (2013), *An Introduction to Statistical Learning: with Applications in* `R`, Springer.
   `http://www-bcf.usc.edu/~gareth/ISL/`

3. Kuan, Chung-Ming (2004), *Statistics: Concepts and Methods* (in Chinese), 2/E, Hwa Tai.
   管中閔 (2004), 統計學: 觀念與方法, 二版, 華泰書局.

4. Chen, Shiu-Sheng (2019), *Probability and Statistical Inference with R* (in Chinese), Tung Hua.
   陳旭昇 (2019), 機率與統計推論, 東華書局.

Course Outline

1. An Introduction to Probability, Statistics & `R`
2. Data Description
3. Probability
4. Univariate Distributions
5. Multivariate Distributions
6. Point Estimation
7. Hypothesis Testing
8. Interval Estimation

Grading

One midterm (**Nov 10, 2021**, 40%), one final (**Jan 5, 2022**, 50%), homework (15%).