

Introduction to Statistics

Lecturer: Prof. Kilkon Ko

Contact: kilkon@gmail.com

Office hours: By appointment

Aims and Goals

This course is specially designed for undergraduate students who want to systematically understand social phenomena through numbers. Although quantitative analysis methods are comprehensive and have been developed over years in different disciplines, the essence of statistics is to answer following two questions: “how to summarize quantitative information” and “how to make inference on the population’s characteristics”.

To tackle with these two questions, we, firstly, learn the statistical theory. This module will cover probability theory, probability and sampling distribution, hypothesis tests of group means, Chi-Square test for categorical variables, and correlation and regression analysis.

During the process of learning these theories, students will have an opportunity to learn how to use statistical software. In this course, students will use EXCEL, SAS, and Python in the class. To help student understand other parts of the world, we will use World Value Survey data. All questions in the midterm and final exam questions should be answered using statistical software.

Textbook

Weiss, N. A. (2005). Introductory statistics. 9th ed. Boston, Addison-Wesley.

Assessment

Class participation (10%), assignment (30%), midterm (30%), final exam (30%)

Course Schedule

Week 1: Introduction: what are scientific methods, quantitative method, and statistics?

- What is a scientific method?
- Difference btw qualitative and quantitative analysis?
- Purpose of statistics

Week 2: Describing and summarizing information: distribution & variation

- SAS Introduction
- Measurement
- Sample and population
- Measures of central tendency
- Measure of variability and dispersion
- Drawing a cross table using EXCEL

Week 3: Visualization of Data

Week 4: Normal Distribution

- Shape of normal distribution
- Characteristics of normal distribution

Week 5: Fundamental concepts for statistical inferences

- Random sampling: treatment, control, randomization, and blinded and double blinded
- Sampling error
- Sampling distribution
- Central limit theorem
- Law of large number

Week 6-7 Confidence Intervals

- CI for the mean with population standard deviation is known
- CI for the mean with population standard deviation is unknown
- Confidence Intervals for proportions
- Margin of error
- EXCEL

Week 8 Midterm Exam

Week 9 Hypothesis testing with a single sample means

- Null and Alternative hypothesis
- Type I, II, and III error
- P-value

Week 10 Hypothesis testing with two sample means

- Independent t-test
- Paired t-test

Week 11: Hypothesis test of categorical variable: Chi-square test

- Association and Independent test

Chapter 11: The Chi-Square Test

Week 12-13: Correlation and Simple Regression

- Assumptions
- Normal equation and BLUE
- Model Fits
- Statistical significance of regression coefficients
- Interpretation of coefficients

Week 14: Multiple Regression

- Assumptions
- Variable Transformation
- Dummy variable
- Partial regression coefficient
- Interpretation of coefficients

Week 15: Final Exam