Parametric Statistics

Sofia Triantafyllou

Fall 2022

E-mail: sof.triantafillou@uoc.gr*
Web: https://polyhedron.math.uoc.gr/2223/moodle/course/view.php?id=3
Class Hours: Tuesday, Thursday 13.15-15.00 (E204)
Office Hours: Tuesdays 15.00-16.30, Thursdays 15.00-16.00
Recitations: TBD
Office: B316

* Please only email me on this email address.

This is a tentative syllabus for the course MEM-262: Parametric Statistics.

Course Description

This is a one-semester course covering the basics of parametric statistics. We will start with a quick introduction to probability theory, and then cover fundamental topics in mathematical statistics such as estimation, hypothesis testing, and convergence.

Required Materials

The course notes are based on the following book:

(DGS) Probability and Statistics, Morris DeGroot and Mark Schervish (International Edition), 2014 (4th edition)

Course notes will be available on the class website. The material can also be found in Greek in the following books:

- Θέματα παραμετρικής στατιστικής συμπερασματολογίας, Σταύρος Κουρουκλής, https://service.eudoxus.gr/search/a/id:59303581/0

Prerequisites

Required preliminary math tools are elementary probability, calculus and basic linear algebra. There are no formal prerequisites for this class, but I would advise that you have passed Calculus I, Calculus II and Probability Theory. Statistics are the flip side of probability. In your probability course, you start with a distribution (say, a normal distribution with a mean μ and a variance σ^2)

and predict features of future observations $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n})$; in parametric statistics we observe the data $\mathbf{x}$ and then try to guess the parameters of the distribution that generated them.

Course Objectives

By the end of the course, you should be familiar with the following statistical concepts:

- Probability basics: Probability is the way to quantify how likely something is to occur. We will talk about experiments, sample spaces, events, independence, conditional probability, and distributions that come up so often they have a name.
- Random Variables and their distributions: Random variables are real-valued functions that link sample spaces and events to data. They model unknown quantities (e.g. the height of a person) in statistical analyses. A random variable takes each one of its possible values with some probability. The collection of these probabilities is the distribution of the random variable. We will discuss random variables and their distributions. We will also discuss some summaries of the distribution, such as expectation and variance, and why they are useful. Finally, we will discuss some mathematical tools for describing the limiting behavior (convergence) of sequences of random variables. This helps us understand what we can expect when we gather more and more data.
- Statistical Inference: Statistical inference is the process of using data to infer (learn) the distribution that generated the data. Sets of distributions that can be described using a finite number of parameters are called parametric models. We will discuss the following types of parametric inference:
  - Point Estimation: Point estimation refers to providing a single "best guess" (estimate) for some quantity of interest, very often a parameter of a parametric model. We will discuss different methods for deriving these guesses: (maximum likelihood, method of moments), as well as their asymptotic properties. We will also discuss how we can choose the best estimator, using statistical decision theory, and how we can quantify uncertainty on the estimated quantities with confidence intervals.
  - Hypothesis testing: Hypothesis testing has to do with trying to decide whether a parameter lies on a specific space of the parameter space or not. For example, we may want to decide if a coin is fair after we have done a number of coin flips. We will discuss how we formalize and test statistical hypotheses, and some specific parametric hypothesis tests.

  We will also discuss the two dominant approaches in statistical inference, frequentist inference and Bayesian inference.

Grading

Grading will be based on the following:

- In-class midterm exam (30%)

- Final Exam (60%)
- Two homework exercises (10%).

Midterm exam will be in-class on October 25th. If there are people who miss the midterm for legitimate reasons, a repeat midterm will be held, date and time decided through voting. **No other arrangements will be made.** September exam will be comprehensive and will count for 90% of the grade. I strongly advise participating in the midterm exam.

Course Syllabus

Week	Content	Chapters in DGS
Week 1: 27/09, 29/09	Introduction/Probabilities	1-2
Week 2: 04/10, 06/10	Random Variables and Distributions	3.1 - 3.4
Week 3: 11/10, 13/10	Random Variables and Distributions	3.5 - 3.9
Week 4: 18/10, 20/10	Expectations & Special Distributions	4-5
Week 5: 25/10, 27/10	Midterm Week	
Week 6: 01/11, 03/11	Large Random Samples	6
Week 7: 08/11, 10/11	Estimation	7.1 - 7.5
Week 8: 15/11, 17/11	Estimation, Sampling Distributions	7.6-7.8, 8.1-8.2
Week 9: 22/11, 24/11	Sampling Distributions	8.3-8.8
Week 10: 29/11, 01/12	Hypothesis Testing	9.1-9.3
Week 11: 06/12, 08/12	Hypothesis Testing	9.5-9.8
Week 12: 13/12, 15/12	Advanced topics	
Week 13: 20/12, 22/12	Review	