

# Parametric Statistics

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete  
Department of Mathematics and Applied Mathematics

September 27, 2022

## Course Info

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

## Course Materials

The course notes are based on the following books

(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 book:

Θέματα παραμετρικής στατιστικής συμπερασματολογίας,  
Σταύρος Κουρουκλής,

<https://service.eudoxus.gr/search/a/id:59303581/0>

# Grading

- ▶ Midterm exams, 30%
- ▶ Final exam, 60%
- ▶ Two Homeworks, 10%

Midterm exam will be in-class during week 5 of the course. 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.

# How to succeed

## 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.

## Study!

- ▶ Two-four hours every week throughout the semester.
- ▶ It is best if you study right after the class.
- ▶ Every week I will post the recitation on the website on Thursday. Try to solve the exercises before you go to the recitation.

# What is statistics

Statistics and probability are concerned with describing uncertainty. In both disciplines, we have some **experiments** that can have **multiple outcomes** that occur with some **probability**.

In your probability course, you start with a distribution (say, a normal distribution with a mean  $\mu$  and a variance  $\sigma^2$ ) and predict features of future observations  $x = (x_1, x_2, \dots, x_n)$ .

# What is statistics

In statistics we observe the data  $x$  and then try to understand the properties of the data-generating process.

In parametric statistics, we assume that the process can be described using a finite number of parameters  $\theta$  and try to identify these parameters.

## Example

We are measuring the height of the students in the class. We assume that they follow a normal distribution. Which are the parameters we need to identify?

# Things we will learn in this course

- ▶ Experiments, Sample Spaces, Events
- ▶ Probability
- ▶ Independence, Conditional Probability
- ▶ Random Variables and their distributions
- ▶ Expectation, variance
- ▶ Large Random Samples
- ▶ Point Estimation
- ▶ Hypothesis testing



## Data Collection

Last year, we performed an ANONYMOUS survey. Students were given an empty card. They wrote their height and gender on the card (answering was optional, they could leave any or both fields blank).

## Let's do some statistical inference

Height	Gender
1.75	F
$\vdots$	$\vdots$
1.68	M

- ▶ Point estimation: Assuming the data follow a Normal  $N(\mu, \sigma^2)$  distribution, what is the "best guess" for  $\mu$  and  $\sigma$ ?
- ▶ Hypothesis testing: Is the mean height of females (significantly) smaller than the height of males? (Do we have the data to test this hypothesis?)

# Statistical inference using R

- ▶ Install R: <https://cran.r-project.org/>
- ▶ Install R-studio:  
<https://www.rstudio.com/products/rstudio/download>
- ▶ Download `class_data_heights.xlsx` from the class website
- ▶ Run the commands in `Rscript0.R` to see the summary of our data.