האוניברסיטה העברית בירושלים THE HEBREW UNIVERSITY OF JERUSALEM



The Hebrew University of Jerusalem

Syllabus

Introduction to Statistics - 52003

Last update 30-09-2021

HU Credits: 4

Degree/Cycle: 1st degree (Bachelor)

Responsible Department: Statistics

Academic year: 0

Semester: 1st Semester

<u>Teaching Languages:</u> Hebrew

Campus: E. Safra

Course/Module Coordinator: Yuval Benjamini

Coordinator Email: yuval.benjamini@mail.huji.ac.il

Coordinator Office Hours: By appointment

Teaching Staff:

Dr. Yuval Benjamini, Mr. Dan Derazne

Course/Module description:

How do we know if the Covid-19 is expanding in Israel? How do we decide if a new vaccine is effective? If our new code made the computer algorithm run faster? In answering these question, we must rely on data that is partial, and make decisions under uncertain conditions.

Statistics deals with methods for making decisions from data, and in mathematically analyzing such methods. In this introductory course we will get to know methods for collecting, summarizing, displaying and making decisions with data.

Course/Module aims:

The aim of the course is to provide a grounding in the aspects of statistics and in particular statistical modelling that are of relevance to practical statistical work.

Learning outcomes - On successful completion of this module, students should be able to:

Basic familiarity with the R statistical programming environment: Input and output, data manipulation, simple simulations, descriptive statistics, statistical inference.

Familiarity with the relations between probability and statistics. The role of probabilistic models in statistics.

Statistical point estimation, confidence intervals, statistical test, correlation, regression. Should be able to choose a mode and carry out inference given data.

Give a verbal description of a problem should be able to propose a method of inference and carry out statistical computations in order to solve the problem.

<u>Attendance requirements(%):</u> None

Teaching arrangement and method of instruction: The course consists of lectures, tirgul, self-work and term project.

There will be a weekly "MATALA" with no grade, and a short exam on moodle with grade.

Each pair of students will submit a term-project consisting of collection and analysis of data. The project is due Jan 20.

Course/Module Content:

a. Principles in data collection. Surveys (population, sample). Experiment (ranodmization, control). Observational studies and the weaknesses. Empirical distribution.

b. Descriptive Statistics.

Summary and graphical description of one or two variables: Location and spread, robustness, order statistics, histogram, scatter, quantile plots. Methods for plotting trend-lines (linear regression, smooth regressions).

c. Estimation and Confidence intervals. Point estimation, bias, variance, estimation methods (moments, maximum likelihood, simulation based). Confidence intervals definition and estimation.

d. Hypothesis testing and examples: location, spread, goodness-of-fit. Nonparameteric methods.

e. Additional topics: Inference for relation between variables. Multiple comparisons.

<u>Required Reading:</u>

Will be uploaded to the moodle.

Additional Reading Material:

1. אלונה רביב ותלמה לויתן, מבוא להסתברות וסטטיסטיקה : הסתברות, עמיחי.

2. אלונה רביב ותלמה לויתן, מבוא להסתברות וסטטיסטיקה : הסקה סטטיסטית, עמיחי.

3. Mathematical Statistics and Data Analysis by John A Rice.

4. Statistics by Freedman Pisani and Purves

<u>Course/Module evaluation:</u> End of year written/oral examination 70 % Presentation 0 % Participation in Tutorials 0 % Project work 0 % Assignments 0 % Reports 0 % Research project 20 % Quizzes 10 % Other 0 %

Additional information:

The course assumes basic familiarity with Probability. We will have an exam in week 2 to test this familiarity. Information can be found on the Moodle website.