The Hebrew University of Jerusalem

Syllabus

Regression and Linear Models (CS) - 52307

Last update 19-03-2025

HU Credits: *4*

Degree/Cycle: *1st degree (Bachelor)*

Responsible Department: *Statistics*

Academic year: *0*

Semester: *2nd Semester*

Teaching Languages: *Hebrew*

Campus: *E. Safra*

Course/Module Coordinator: *Or Zuk*

Coordinator Email: *or.zuk@mail.huji.ac.il*

Coordinator Office Hours: *Monday 10-11*

Teaching Staff:

Dr. Or Zuk,
Mr. Hadar Sharvit

Course/Module description:
Linear regression.

Course/Module aims:
To learn the theoretical basis of linear regression, as well as how to apply it in analyzing data sets.

Learning outcomes - On successful completion of this module, students should be able to:
Define a linear regression model.
Compute the corresponding estimates and test statistics.
Construct a model appropriate for a given problem and set of data.
Apply the estimates and test statistics to analyze the data.
Perform the calculations using a computer.
Understand the geometrical interpretation of regression.

Attendance requirements(%):
0

Teaching arrangement and method of instruction: The instructor will teach the theory and write on the board. In addition, he will display many examples with computer output (data sets, results of statistical analyses, etc.)

Exercises and examples will be uploaded to the course's website

Course/Module Content:
Simple linear regression: Examples, formulation and interpretation of the model, estimation using ordinary least squares.

Multiple linear regression: Examples, formulation and interpretation of the model - defining explanatory variables, transformations, interactions. Estmation using ordinary least squares and maximum likelihood, multiple correlation coefficient, multicollinearity

The geometric interpretation of regression: subspaces, orthogonality, projections, Pythagorean Theorem. Distribution of the estimates and statistical inference. Verifying the appropriateness of the model - examining residuals, measures of influence, partial residuals, variance-stabilizing transformations.

Building models: Stepwise methods for choosing subsets of explanatory variables, multicollinearity, definitions of "good equations", the Cp statistic.

Logistic regression for binary outcomes

Analysis of Variance (ANOVA), 1-way and 2-way

Models with dependent errors: Examples, formulation of a model, the likelihood function, estimation using weighted least-squares.

Required Reading:
Required reading material will be posted on the course website over the course of the semester

Additional Reading Material:
Additional recommended reading material will be posted on the course website over the course of the semester.

References:
Draper, N. and Smith, H. (1998). Applied Regression Analysis.
Daniel, C. and Wood, F. (1971). Fitting Equations to Data.
Weisberg, S. (1980). Applied Linear Regression.

Grading Scheme:
Written / Oral / Practical Exam 80 %
Submission assignments during the semester: Exercises / Essays / Audits / Reports / Forum / Simulation / others 20 %

Additional information:
Course Requirements

1. Final exam (in standard format)

2. Exercises: A few exercises for self practicing will be given during the semester, in addition to a big exercise (avodat bit) that must be submitted and comprises 20% of the course grade.
The exercises are an integral part of the course, and investing effort on them is essential to proper learning of the material. It is possible that some of the exercise questions will appear on the final exam (with some changes).

*To pass the course, a passing grade (60) in the final exam is required. The exercises will be taken into account in the final course grade only for those who get a passing grade on the final exam. For those who do not pass the final exam, the final course grade will be the *minimum* of the weighted course grade and the final exam grade.*

Remark:
During the course we will be using the statistical package R. The package can be downloaded free from the following site:
http://cloud.r-project.org/