



Scuola Superiore  
Sant'Anna



## **Applied Statistics for Social Science**

Chiara Seghieri, Laboratorio MeS, Istituto di Management, Scuola Superiore Sant'Anna. [c.seghieri@sss sup.it](mailto:c.seghieri@sss sup.it)

### **Course Description**

The course intends to provide a systematic introduction to quantitative approaches to data collection and analysis in the social sciences. Emphasis will be placed on the development of statistical concepts and statistical computing skills, rather than mathematical details.

An introduction to Stata, software for statistical analysis is also provided along the course. This will be useful for the Stata course by Prof. Nicola Orsini.

Students will have the opportunity to develop a critical understanding of the use of statistics in contemporary social science, to think about their own data, to think of ways to assess the data's appropriateness, extracting the most information using the most common statistical methods and to critically review papers that make use of statistics.

### **Lectures Topics**

Elements of sample surveys, experimental design and observational studies; methods of parameter estimation and hypothesis testing in one- and two-sample problems; linear regression model with one or more predictors; logit and probit models, introduction to multivariate analysis (cluster and factor analysis).

Upon successful completion of this course, the students should be able to describe a sample and know which statistics are appropriate for measuring center and spread, understand randomness, sampling techniques, and experiments, conduct hypothesis tests for proportions and means of one and two samples, apply the classical, p-value, and confidence interval approaches to hypothesis testing, regression analysis interpretation, presentation and diagnostics.

## **Prerequisites**

The course will assume knowledge of descriptive statistics (Frequency Distributions and their Graphic Representations, Measures of Central Tendency, Measures of Dispersion). These topics will NOT be covered during the course and students will therefore have to study them during Summer 2013. Please write to Chiara Seghieri for any doubt about topics which will be taken for granted.

## **Assignments:**

During the course, weekly assignments will be given, which are combination of readings and computer exercise using Excel/Stata.

## **Suggested Books**

It is strongly suggested/required that you choose at least one Statistics textbook to study before the course. Basic statistic concepts are taken for granted!. Here some example of text books:

- Fox, John. 1997. Applied Regression Analysis, Linear Models, and Related Methods. Thousand Oaks, CA: Sage
- Freedman, David; Robert Pisani; and Roger Purves. 1998. Statistics. 3rd Edition. New York: Norton.
- Agresti, Alan and Finlay, Barbara. 1997. Statistical Methods for the Social Sciences Upper Saddle River, NJ: Prentice Hall.

## **Tentative Course Schedule, October-November 2013:**

Descriptive statistics: class exercise discussion 31/10/2013

Probability and Sampling Distributions: 05/11/2013 - 10:00-12:00

*Best Practices in surveying process, Questionnaire design. Dott.ssa Michela Natilli*

Probability and Sampling Distributions B 07/11/2013 - 10:00-12:00

*Sampling techniques, sample size calculation. Dott.ssa Michela Natilli*

Inference: Estimators and Tests Part A: 12/11/2013 - 10:00-12:00

Inference: Estimators and Tests Part B: 14/11/2013 - 10:00-12:00

Linear Model: Bivariate and Multiple Regression 19/11/2013 - 10:00-12:00

*Relation between two variables, Regression coefficient estimation, testing of the parameters. Dott.ssa Michela Natilli*

Linear Model: Interpretation, Hypothesis Testing, Diagnostics A 21/11/2013 - 10:00-12:00

*Relation among  $n$  variables, regression coefficients estimation, testing of the parameters. Dott.ssa Michela Natilli*

Linear Model: Interpretation, Hypothesis Testing, Diagnostics B 26/11/2013 - 10:00-12:00

*Model diagnosis and Regression diagnostics. Dott.ssa Michela Natilli*

GLM: Logit/Probit Part A 28/11/2013 - 10:00-12:00

GLM: Logit/Probit Part B 03/12/2013 - 10:00-12:00

Multivariate analysis: Cluster Analysis 05/12/2013 - 10:00-12:00

Multivariate analysis: Factor analysis 10/12/2013 - 10:00-12:00