



Scuola Superiore  
Sant'Anna



Phd in Management

## **Language of Stata statistical software**

Nicola Orsini, Karolinska Institutet

Email: [nicola.orsini@ki.se](mailto:nicola.orsini@ki.se)

Web: <https://ki.se/en/people/nicors>

**Dates:** June 12-15, 2018

### **Short description of the course:**

This course is designed to introduce doctoral students to the language of Stata statistical software in order to analyze, interpret, and present both numerically and graphically the findings.

As motivating and instructive example both lectures and labs will be based on the analysis of a real study published in one of the most prestigious scientific journals in science (ISI Impact Factor=54). The course emphasizes how fluent use of Stata language can facilitate research and academic publishing.

Specific topics are: assessing the functional (i.e. J-shaped, U-shaped, linear) relation between a continuous independent variable and the response variable using flexible tools (i.e. splines), plotting changes in the predicted response as function of covariates with confidence intervals, building a multivariable model, assessing and interpreting interactions between explanatory variables, analyzing both continuous and binary outcomes.

### **Student commitment**

Upon enrolling into this course, students commit to come to class both for lectures and labs. Attendance is compulsory. Students need to make sure they have Stata installed on their laptops.

## Schedule

Each session (morning and afternoon) includes three parts: lecture, exercise, and review to the exercise.

<b>Date</b>	<b>Main topic</b>	<b>Description</b>	<b>Schedule</b>
June 12, 2018	The Basics	Univariate analysis of categorical and continuous variables	09:00-13:00 14:00-17:00
June 13, 2018	Models for continuous outcomes	Linear regression Predicted Mean Differences Testing hypothesis	09:00-13:00 14:00-17:00
June 14, 2018	Models for binary outcomes	Logistic/Probit regression Predicted responses and their changes	09:00-13:00 14:00-17:00
June 15, 2018	Advanced modelling strategies	Splines, interaction between predictors Missing data	09:00-13:00 14:00-17:00

## Course material

Lecture notes, links, further readings, datasets, exercises, and solutions will be available at <http://nicolaorsini.altervista.org/sanna/is.htm>