Introduction to

Probability, Statistics and Econometrics

EUI September 2016

Course outline: version of April 18, 2016

Instructors

Part 1: Filip Lazaric (filip.lazaric@eui.eu)

Part 2: Andrea Ichino (andrea.ichino@eui.eu)

Exercise classes: Madina Kurmangaliyeva (madina.kurmangaliyeva@eui.eu)

General Introduction

This course is an introduction to probability, statistics and econometrics with emphasis, given the time constraint, on the topics that students need to know and master well for the sequence of more advanced econometric courses that will follow in the PhD program at the EUI.

A knowledge of calculus and of some elementary statistics and matrix algebra is required. Students who have doubts on whether they satisfy these pre-requisites should talk to the instructors to arrange special training.

Lectures will refer to the following textbooks that are available in the library at VSP and to classnotes that will be downloadable at an URL that will be indicated later:

- Richard J. Larsen and Morris L. Marx. *An introduction to mathematical statistics and its applications.* Prentice Hall, Fifth Edition, 2012.
- George Casella and Roger L. Berger. *Statistical Inference.* Thomson, Second Edition, 2002.
- Adrian Papamarcou. *Notes on Probability Theory* (with references to "Probability and Measure" by Patrick Billingsley). 2006.
- Jeffrey Wooldridge, *Introductory Econometrics. A Modern Appproach.* South Western Cengage Learning, 2009
- Joshua Angrist and Jorn-Steffen Pischke. *Mostly Harmless Econometrics. An Empiricist's Companion.* Princeton University Press, 2013.
- Lecture notes by the instructors.

1 Part 1: Probability and statistics

The main goal of this Part is to give an introduction to the axiomatic foundations of probability theory and to the basics of univariate and bivariate statistics. The Part will consist of two parts of approximately equal size. The first part will focus on probability theory and will provide a mathematical structure for measures of uncertainty. In the second part the concept of random variables will be introduced and its properties will be discussed. Both parts will contain examples illustrating the basic concepts.

There will be eight lectures and two exercise classes in this part.

Tentative Schedule

Lecture 1

Random Experiment, Event Space, σ-algebra, Generated σ-algebra, Borel-algebra

Larsen and Marx, chapter 2. Casella and Berger, chapter 1.Papamarcou, p. 6-18.

Lecture 2

Approaches to Probability, Measures, Probability set functions

Larsen and Marx, chapter 2. Casella and Berger, chapter 1. Papamarcou, p. 18-26.

Lecture 3

Conditional Probability, Law of Total Probability, Bayes Theorem, Independence, Basic Elements in Counting and Combinatorics

Larsen and Marx, chapter 2. Casella and Berger, chapter 1.

Lecture 4

Measurable Functions and Random Variables, Distribution Functions of Random Variables

Larsen and Marx, chapter 3. Casella and Berger, chapter 1.5 - 1.6. Papamarcou, p. 26 - 42

Lecture 5

Change of variables, Moments of Random Variables, Moment generating function, Markov-, Chebychev- and Jensen's inequality

Larsen and Marx, chapter 3. Casella and Berger, chapter 2. Papamarcou, p. 52 - 64.

Lecture 6

Some standard distributions and their properties

Larsen and Marx, chapter 4. Casella and Berger, chapter 3.

Lecture 7

Bivariate Distributions: Joint CDFs, Marginal Distributions, Conditional Distributions

Larsen and Marx, chapter 3. Casella and Berger, chapter 4.

Lecture 8

Covariance and Correlation, Conditional Expectation,Independence

Larsen and Marx, chapter 3. Casella and Berger, chapter 4.

Exercise class 1

Deadline for first problem set

Exercise Class 2

Deadline for second problem set

2 Part 2: Statistics and Econometrics

The main goal of this Part is to give an introduction to the basic tools that an econometrician needs: the most popular estimation methods; inference and hypothesis testing; asymptotics; simple and multiple regression. Approximately two lectures will be devoted to each of the first three topics while regression will absorb four lectures.

Therefore, there will be ten lectures in this Part plus four exercise classes. Examples and applications will be used in the lectures and particularly in the exercise classes, to illustrate the theoretical content of the course.

Tentative Schedule

Lecture 1

Introduction: what is econometrics about; the tool-box of econometrics; the econometrics sequence at the EUI; Content of this course.
Estimation: Estimators and estimates; the Method of maximum Likelihood; the Method of Moments.

Larsen and Marx, chapter 5. Casella and Berger, chapter 7 . Lecture notes.

Lecture 2

Estimation: Finite sample properties of estimators; Unbiasedness, Efficiency,, Sufficiency, Minimum variance estimators; The Cramer-Rao Lower Bound, Invariance.

Larsen and Marx, chapter 5. Casella and Berger, chapter 7 and chapter 5. Lecture notes.

Lecture 3

Estimation: Asymptotic properties of estimators; Asymptotic Unbiasedness, Asymnptotic Efficiency, Consistency; Asymptotic Normality
Basic asymptotics: concepts of convergence; Law of Large Numbers; Central Limit theorem; Continuous Mapping Theorem, Slutzky Theorem and Delta Method.

Larsen and Marx, chapter 5. Casella and Berger, chapter 7 and chapter 5. Lecture notes.

Lecture 4

Simple regression: The Conditional Expectation Function; The Population Regression Function; The Sample Regression Function; OLS,

Method of Moments and Maximum Likelihood estimation of a regression; Algebraic and geometric properties of the OLS-MM estimators.

Angrist and Pischke chapter 1, 2 and 3. Wooldridge part 1. Lecture notes.

Lecture 5

Simple regression: Goodness of fit and the R-Squared; Statistical Properties of the OLS-MM estimator; The Gauss-Markov Theorem'.

Angrist and Pischke chapter 1, 2 and 3. Wooldridge part 1. Lecture notes.

Lecture 6

Simple regression: Causality and Regression.

Angrist and Pischke chapter 1, 2 and 3. Lecture notes.

Lecture 7

Multiple regression: The Conditional Independence Assumption; Interpretation of the partial Multiple Regression Coefficient; Multiple Regression in matrix notation; Omitted variable bias and inclusion of irrelevant regressors.

Angrist and Pischke chapter 1, 2 and 3. Wooldridge part 1. Lecture notes.

Lecture 8

Multiple regression: The Gauss-Markov Theorem and Multiple Regression; "Partialling out" and the interpretation of coefficients; Good and bad habits concerning control variables;

Angrist and Pischke chapter 1, 2 and 3. Wooldridge part 1. Lecture notes.

Lecture 9

Inference and Hypothesis testing: what is a statistical test and how it is constructed; The decision rule; Type I and type II errors; Power of a test.

Larsen and Marx, chapters 6 and 9. Casella and Berger, chapter 8. Lecture notes.

Lecture 10

Inference and Hypothesis testing: finite sample and asymptotic tests in the context of a regression model.

Larsen and Marx, chapters 6 and 9. Casella and Berger, chapter 8. Lecture notes

Exercise classes will take place on Thursdays

1. **Exercise Class**
   Deadline for first problem set
2. **Exercise Class**
   Deadline for second problem set
3. **Exercise Class**
   Deadline for third problem set
4. **Exercise Class**
   Deadline for fourth problem set

3 Final exam and Grading

The exam for both Parts will be in the exam week

The grade will be based on:

- six problem sets (two for Part 1 and four for Part 2) which will count for 20% of the final grade;
- a final classroom exam in which questions on the material of Part 1 will count for 20% of the final grade and questions on the material of Part 2 will count for 60% of the final grade.

The questions in the exam will be of the following types:

- eight "true and false questions"; these are statements that the student will have to comment briefly, deciding whether they are true of false. What counts is not the true/false answer per se (in most cases both could be defended) but the line of reasoning chosen by the student to support her/his preferred answer(s).
- one exercise based on the material of Part 1;
- one longer exercise based on the material of Part 2.