Course guide
200604 - IEA - Advanced Statistical Inference

Unit in charge: School of Mathematics and Statistics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.
1004 - UB - (ENG)Universitat de Barcelona.

Degree: MASTER'S DEGREE IN STATISTICS AND OPERATIONS RESEARCH (Syllabus 2013). (Optional subject).

Academic year: 2022   ECTS Credits: 5.0   Languages: English

LECTURER

Coordinating lecturer: GUADALUPE GÓMEZ MELIS

Others: Primer quadrimestre:
GUADALUPE GÓMEZ MELIS - A
ÁLEX SÁNCHEZ PLA - A

PRIOR SKILLS

Advanced Statistical Inference is mandatory for all graduate students in statistics or mathematics.
To follow and take advantage of this course the statistical knowledge required is that of an undergraduate-level in statistics or mathematics.

* Basic mathematical analysis skills required: integration of functions of one or two variables, derivation, optimization of a function of one or two variables.
* Basic probability skills required: the most common parametric distributions, properties of a normal distribution, the law of large numbers and the central limit theorem.
* Basic statistical inference skills required: using the likelihood function for simple random sampling (independent identically distributed data), inference in the case of normality, estimation of maximum likelihood for parametric models with only one parameter and simple random sampling.

Chapter 1 in Wood's "Core Statistics" and Chapter 1 in Gomez and Delicado's "Inference and Decision" include all the concepts and results that are assumed to be known. Students will be required to review, achieve and internalize them before beginning the course. A list of exercicis non-assessable but compulsory will be held in the second week of classes.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
3. CE-1. Ability to design and manage the collection of information and coding, handling, storing and processing it.
4. CE-4. Ability to use different inference procedures to answer questions, identifying the properties of different estimation methods and their advantages and disadvantages, tailored to a specific situation and a specific context.
5. CE-6. Ability to use appropriate software to perform the necessary calculations in solving a problem.
6. CE-9. Ability to implement statistical and operations research algorithms.

Transversal:
2. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.
TEACHING METHODOLOGY

Conceptual sessions of 1.5 hours

The sessions present the subject material. The teacher might use the computer to present the contents. Ideas and concepts are emphasized and a detailed look is given at those proofs with an added pedagogical value.

Chapters 2, 4 and 5 in the book "Core Statistics" from Simon Wood will be followed. Most of the material can be as well read in Chapters 1 to 5 of Gomez and Delicado (although in different order), these notes can be downloaded from the Intranet. Supplementary materials will be provided for specific topics.

Problem sessions of 1.5 hours.

Problems will be posted on the intranet and will be discussed in class. Students must come to class having prepared the problems and having thought about how to solve them. The professor will solve the problems and discuss with students their questions or other solutions. The solution of these problems will be posted after the corresponding session on the intranet.

Statistical Laboratory

R programs will serve to illustrate concepts, to complement the theoretical developments showing how statistical computing is an important tool in statistical inference.

- Some exercises will be proposed in line with those discussed in class, to strengthen the concepts.

LEARNING OBJECTIVES OF THE SUBJECT

The Advanced Statistical Inference course provides a theoretical and applied basis for the fundamentals of Statistics. Its main objective is to train students to think in statistical terms in order to conduct a thorough professional habit. Also intended as a formative seed for the consolidation of young researchers in this area of science and technology while equipping students with the resources to continue their training and making them capable to read papers published in journals of statistics.

After completing the course the student:
* has learnt about the different principles governing the reduction of a dataset and the different philosophies that may arise to solve a problem.
* knows the methods based on the empirical distribution function and the likelihood function and know when and why how apply each one
* understands that the frequentist and Bayesian philosophy are two ways to approach a problem, not necessarily conflicting and sometimes complementary.
* be familiar with modern resampling techniques and view them as a formal/computational approximation well suited for use in situations where direct calculations are too complex or not available
* know to write down the likelihood function in different situations and learn different techniques to maximize it.

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours small group</td>
<td>15,0</td>
<td>12.00</td>
</tr>
<tr>
<td>Hours large group</td>
<td>30,0</td>
<td>24.00</td>
</tr>
<tr>
<td>Self study</td>
<td>80,0</td>
<td>64.00</td>
</tr>
</tbody>
</table>

**Total learning time:** 125 h
## CONTENTS

### 1. The foundations of Statistical Inference

**Description:**
- Preliminaries, notation and examples
- Inferential questions. Walking through point estimation, hypothesis testing and interval estimation
- The frequentist approach: point estimation, finite sample properties, Cramer-Rao bound, Hypothesis testing, Interval estimation, Model checking and model comparison
- The Bayesian approach: a brief discussion

**Full-or-part-time:** 52h 50m  
Theory classes: 15h  
Practical classes: 4h 30m  
Self study : 33h 20m

### 2. The empirical distribution function. Theory and numerical approaches

**Description:**
- The empirical distribution function. Glivenko-Cantelli Theorem.
- Principle of substitution. The method of moments.
- Introduction to bootstrap.
- Large sample properties: Delta method and consistency

**Full-or-part-time:** 32h  
Theory classes: 9h  
Practical classes: 3h  
Self study : 20h

### 3. Maximum Likelihood Estimation. Theory and numerical approaches

**Description:**
- Likelihood, log likelihood and score functions
- Fisher information matrix, Cramer-Rao bound and UMVUE
- Large sample properties of MLE. Consistency and Asymptotic Normality
- Generalised Likelihood Ratio Statistic
- AIC information criterion
- Numerical approaches
- EM algorithm

**Full-or-part-time:** 40h 10m  
Theory classes: 9h  
Practical classes: 4h 30m  
Self study : 26h 40m
GRADING SYSTEM

The assessment of Chapter 1 relies on a partial exam (EP). The partial exam (EP) will contain a theoretical part and some problems. For the assessment of Chapters 2 and 3 there will be 2 assignments of problems/practices with R (PRA) and a final exam (EF). The delivery of problems will be done at most in groups of two. The final exam (EF) consists of problem solving.

The final mark (N) is obtained from the marks of the delivered exercises and the marks of the partial and final exams according to the expression:

\[ N = 0.25 \times \text{PRA} + 0.25 \times \text{EP} + 0.5 \times \text{EF}. \]

BIBLIOGRAPHY

Basic:

Complementary: