Course guide
200604 - IEA - Advanced Statistical Inference

Last modified: 14/05/2024

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: 2024 **ECTS Credits:** 5.0 **Languages:** English

LECTURER

Coordinating lecturer: GUADALUPE GÓMEZ MELIS

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

PRIOR SKILLS

Advanced Statistical Inference is mandatory for all graduate students in statistics, mathematics, data science, physics and all those who have an intermediate background in statistical inference
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 uni parametric models.

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.

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

1.5 hour theory sessions

These are sessions to present the main material of the subject. The professor uses slides and the blackboard to present the concepts trying to emphasize the main ideas and the rationale behind the arguments. The professor will look in detail at those proofs that are pedagogically creative and formative.

Chapters 2, 4 and 5 of the book "Core Statistics" by Simon Wood will be followed, as well as the material
in the notes of Gómez and Delicado that can be downloaded from the Intranet.
Other complementary materials for specific topics will be provided.

Sessions of problems of 1.5h.

Tthe problems that will be discussed at each session will be posted on the intranet one week before.
The students should read and solve the problems in the list before class.
The professor will solve some of the problems and discuss doubts or other solutions with the students.
The solution to some of these problems is posted after the corresponding session on the intranet.

Statistics Laboratories
R code will be provided, in the form of scripts or "notebooks", to illustrate various concepts, as well as for the application of methods such as Bootstrap or numerical approximations complementing the developments shown in theory.

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 to conduct a thorough professional career. 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, why and how apply each one.
* 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
* will know how to set up the likelihood function in various situations and will know different techniques, analytical or numerical, to maximize it.

STUDY LOAD

Type	Hours	Percentage
Hours small group	15,0	12.00
Self study	80,0	64.00
Hours large group	30,0	24.00

Total learning time: 125 h

CONTENTS

1.The foundations of Statistical Inference

Description:
- Preliminars, 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.
- 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.
- Resampling methods (bootstrap).
- Large sample properties: consistency, asymptotic normality and Delta method.

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

For the evaluation of each topic there will be an exam. The exam for topic 1 (EX1) will be held 2-3 weeks after finishing topic 1. The exams for topics 2 and 3 (EX2 and EX3) will be taken on the day assigned for the final exam in January. Those who have failed the topic 1 exam (or choose to improve their mark) will have the option to retake the exam in January. The exams will consist of a theoretical part and some problems.
During the course, 4 questionnaires (Q) will be passed.

The mark by continuous avaluation, NC, is obtained according to the expression:
NC= 0.30 * EX1+0.50*EX2+0.20*Q
The mark by unique avaluation, done during January's exam, is obtained according to the expression:
NU= 0.40 * EX1+0.60*EX2
The final mark of the subject (NF) is obtained according to the expression:
NF= max{NC, NU}

BIBLIOGRAPHY

Basic:
- Wood, Simon N. Core statistics. Cambridge [etc.]: Cambridge University Press, 2015. ISBN 9781107071056.
- Gómez Melis, G.; Delicado, P. Inferencia y decisión (apunts). 2003.
- Olive, David J. Statistical theory and inference. Cham: Springer, 2014. ISBN 9783319049717.
- Casella, G.; Berger, Roger L. Statistical inference. Pacific Grove Duxbury, 2002. ISBN 0534243126.
- Cox, D.R. Principles of statistical inference. Cambridge Univ Press, 2006. ISBN 9780521685672.
- Wasserman, Larry. All of statistics : a concise course in statistical inference [on line]. Pittsburgh: Springer, 2004 [Consultation: 04/07/2023]. Available on: https://rss-onlinelibrary-wiley-com.recursos.biblioteca.upc.edu/doi/full/10.1111/j.1467-985X.2004.00347_18.x. ISBN 9781441923226.
- Held, Leonhard; Sabanés Bové, Daniel. Likelihood and bayesian inference. Second. Springer, 2020. ISBN 978662607947.

Complementary:
- Boos, D.D.; Stefanski, L.A. Essential statistical inference : theory and methods [on line]. Springer, 2013 [Consultation: 04/07/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/978-1-4614-4818-1. ISBN 9781461448174.
- Chihara, L. ; Hesterberg, T. Mathematical statistics with resampling and R. Wiley, 2011. ISBN 9781118029855.
- Ruiz-Maya Pérez, L. ; Martin Pliego, F.J. Estadística. II, inferencia. 2ª ed. Madrid: Alfa Centauro, 2001. ISBN 8472881962.
- Shao, Jun. Mathematical statistics [on line]. 2nd ed. Springer Texts in Statistics, 2003 [Consultation: 04/07/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b98900.
- Young, G.A.; Smith, R.L. Essentials of statistical inference. Cambridge University Press, 2010. ISBN 9780521548663.
- Cuadras, C. M. Problemas de probabilidades y estadística. Vol 2: Inferencia [on line]. Publicacions de la Universitat de Barcelona, 2016 [Consultation: 04/07/2023]. Available on: https://lectura-unebook-es.recursos.biblioteca.upc.edu/viewer/9788491685203/1. ISBN 9788491685203.