# UNIVERSITÀ DEGLI STUDI DI PALERMO

| | | | |
|---|---|---|---|
| **DEPARTMENT** | Scienze Economiche, Aziendali e Statistiche | | |
| **ACADEMIC YEAR** | 2020/2021 | | |
| **MASTER'S DEGREE (MSC)** | STATISTICS AND DATA SCIENCE | | |
| **INTEGRATED COURSE** | ADVANCED STATISTICAL METHODS - INTEGRATED COURSE | | |
| **CODE** | 21226 | | |
| **MODULES** | Yes | | |
| **NUMBER OF MODULES** | 2 | | |
| **SCIENTIFIC SECTOR(S)** | SECS-S/01 | | |
| **HEAD PROFESSOR(S)** | CHIODI MARCELLO | Professore Ordinario | Univ. di PALERMO |
| **OTHER PROFESSOR(S)** | ABBRUZZO ANTONINO<br>CHIODI MARCELLO | Professore Associato<br>Professore Ordinario | Univ. di PALERMO<br>Univ. di PALERMO |
| **CREDITS** | 9 | | |
| **PROPAEDEUTICAL SUBJECTS** | | | |
| **MUTUALIZATION** | | | |
| **YEAR** | 2 | | |
| **TERM (SEMESTER)** | 1° semester | | |
| **ATTENDANCE** | Not mandatory | | |
| **EVALUATION** | Out of 30 | | |
| **TEACHER OFFICE HOURS** | **ABBRUZZO ANTONINO**<br>Monday 15:00 17:00 DSEAS secondo piano stanza 222<br>**CHIODI MARCELLO**<br>Tuesday 15:00 17:00 stanza del docente (edificio 13); eccezionalmente su teams<br>Friday 12:00 13:00 stanza del docente (edificio 13); eccezionalmente su teams | | |

**DOCENTE:** Prof. MARCELLO CHIODI

| | |
|---|---|
| **PREREQUISITES** | Elements of probability and inferential statistics. Hypothesis test theory. Generalized linear models. |
| **LEARNING OUTCOMES** | Knowledge of basic methods of Bayesian and nonparametric statistics. Acquisition of language and terminology of the discipline. Understanding of derivations, theoretical properties and relations among the presented methods.<br><br>Ability to deal with concrete problems with the methods acquired during the lectures. Ability to use the statistical environment R to apply the skills students have acquired during the lectures and to check via simulation the theoretical results.<br><br>Being able to critically understand the characteristics, potential, and limits of Bayesian, nonparametric and semiparametric statistics. Being able to frame a specific problem in nonparametric terms.<br><br>Being able to discuss the characteristics of a given problem. Being able to use the statistical terminology and the formalization of the problems in writing.<br><br>Being able to see the scientific literature; ability to learn the patterns of extensions studied in class; learning the ability of specialized statistical software also different from that used in the classroom. |
| **ASSESSMENT METHODS** | The final examination will consist of a discussion, and it depends on the fact that the student had passed the written test.<br><br>Written test<br>The written test strives to establish the knowledge and abilities possessed by the student. The written test will cover bayesian, non-parametric and semi-parametric concepts, acquired during the course and it will be carried out with the support of the statistical software R. The test lasts a maximum of 3 hours.<br><br>The sufficiency (equivalent to a score of 18 on a scale of 18 to 30), which is necessary to pass the test, is reached if the student shows an adequate use of the terms relating to the concepts in question, and i) in case of practical question, by identifying the appropriate statistical methodology even if it is spoiled by the mere computation error; ii) in the case of a theoretical question, in the<br>consistency of the answer, although not completely exhaustive of the topic.<br><br>Oral examination<br>The oral test is intended to dig up the topic of the written test and to evaluate the knowledge of the students on the subject. This will consist of at least two questions aimed at graduate better knowledge and abilities possessed by the student, and its ability to provide it with a suitable statistical language. The sufficiency of the oral test will be reached when the student has knowledge and understanding of the topics at least in the general terms (definition of the concepts). The more, however, the examination has brilliantly passed the written test and has shown, in the oral test, their capacities, as well as the status of statistical language and the connection with the other subjects, much more the evaluation will be positive.<br><br>FINAL ASSESSMENT<br>The final evaluation of the examination will take into account two aspects: i) mastery of the topics; ii) the property of statistical language, assessed on the whole of the written and oral test. The teacher will also have the opportunity to take into account the context factors of the exam (such as active participation during the lessons and exercises or the presence of some disabilities) for the purpose of determining the outcome of the test.<br><br>Both tests (written and oral) are evaluated in thirtyth and are considered to be passed with a minimum vote of 18/30. The final score is given by the simple arithmetic mean of the two tests. |
| **TEACHING METHODS** | The course will be divided into lectures and practicals. All the theoretical arguments developed during the lectures will be addressed in terms of applications, by means of computer-statistical practice, with the use of the program environment R. |

## MODULE
## BAYESIAN STATISTICS

*Prof. ANTONINO ABBRUZZO*

**SUGGESTED BIBLIOGRAPHY**

A first course in Bayesian statistical methods, Hoff, Peter D, 2009, Springer Science & Business Media. Capitoli da 5 a 11.

Bayesian Modelling using WinBUGS, Ntzoufras Ioannis, 2009, Wiley. Capitoli da 3 a 8.

Appunti del docente.

| | |
|---|---|
| **AMBIT** | 21031-Attività formative affini o integrative |
| **INDIVIDUAL STUDY (Hrs)** | 54 |
| **COURSE ACTIVITY (Hrs)** | 21 |

**EDUCATIONAL OBJECTIVES OF THE MODULE**

The course guides the student to the knowledge of the basic methods of Bayesian statistics and the acquisition of the ability to apply these methodologies to real datasets. Students should be able: to understand both the positive and negative aspects of Bayesian statistics with respect to the classical ones; and to use these techniques to investigate real datasets.

## SYLLABUS

| Hrs | Frontal teaching |
|---|---|
| 3 | Bayesian inference with Gibbs sampler |
| 3 | Bayesian Inference with the Metropolis-Hastings |
| 3 | Approximate Bayesian Inference methods |
| 3 | Integrated nested Laplace approximation for Bayesian inference |

| Hrs | Practice |
|---|---|
| 2 | Application of Gibbs sampler |
| 3 | Application of the Metropolist-Hastings algorithms |
| 2 | Applicazione dell'algoritmo approssimato Bayesiano |
| 2 | Application of INLA |

## MODULE
## NON PARAMETRIC STATISTICS

*Prof. MARCELLO CHIODI*

**SUGGESTED BIBLIOGRAPHY**

Density Estimation for Statistics and Data Analysis - B.W. Silverman. Statistics and Applied Probability, London: Chapman and Hall, 1986. (Capitoli: 1, 3, 4, 5 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Applied smoothing techniques for data analysis: The kernel approach with splus illustrations. - ADRIAN W. BOWMAN, ADELCHI AZZALINI. CLARENDON PRESS • OXFORD 1997. (Capitoli 1, 2, 3, 4, 5, 6 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Generalized additive models. An Introduction with R. S.N. Wood. Statistical Science: Chapman & Hall, 2006. (Capitoli: 2, 3, 4, 5 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Materiale didattico (dispense e lucidi) forniti dal docente.

| | |
|---|---|
| **AMBIT** | 21031-Attività formative affini o integrative |
| **INDIVIDUAL STUDY (Hrs)** | 108 |
| **COURSE ACTIVITY (Hrs)** | 42 |

**EDUCATIONAL OBJECTIVES OF THE MODULE**

The course aims to guide the student to the knowledge of the basic methods of non parametric statistics and the acquisition of the ability to apply these methodologies to real datasets. Students should be able to understand both the positive and negative aspects of nonparametric and semiparametric statistics with respect to the parametric ones; use these techniques to investigate real datasets.

## SYLLABUS

| Hrs | Frontal teaching |
|---|---|
| 2 | Introduction to nonparametric statistics. Nonparametric hypothesis test, Functional Hypothesis, non parametric estimation |
| 6 | Univariate kernel density: Histogram. Kernel density and their statistical properties. Choice of h. Variable kernel and their properties. Cross-Validation and k-fold cross-validation. MISE minimization. Theoretical properties and simulations |
| 8 | Multivariate kernel density: Multivariate kernel estimator and choice of H. Properties of multivariate estimator: AMISE.Discriminant analysis and cluster analysis.Kth nearest neighbour |
| 8 | Nonparametric regression: Local Polynomial regression, Cubic splines, and penalized spline regression. Quantile regression. Augmented objective function. Choice of the smoothing parameter. Generalized additive models and methods to obtain the estimation of the regression parameters. Bayesian non-parametric techniques |

| Hrs | Practice |
|---|---|
| 3 | Nonparametric hypothesis test and their application to real dataset. |
| 3 | Univariate kernel density estimators and their applications to real datasets. exercises with R. uso of the didactical software MLANP |
| 6 | Multivariate kernel density estimators and their applications to real datasets |
| 6 | Semiparametric and nonparametric regressions and their applications to real datasets. |