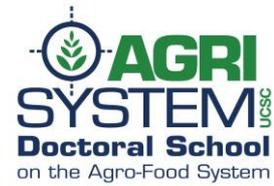




UNIVERSITÀ
CATTOLICA
del Sacro Cuore



ADVANCED METHODS FOR AGRI-FOOD DATA ANALYSIS

Prof. Riccardo Negrini

*Faculty of Agriculture, Food and Environmental Sciences
Università Cattolica del Sacro Cuore – Piacenza*

GENERAL COURSE OBJECTIVES

Data mining and predictive models are at the heart of successful research and many other technology-based solutions to key topics in agri-food and agri-business.

The course “Advanced Methods for Agri-food Data Analysis” is designed as a logical follow up of the course “Introduction to applied statistics for agri-food data”

The statistical methods covered by the course are useful for many types of questions that relate to multiple variables and/or multiple groups. Learning how to effectively use data and statistical methods to make evidence based decisions is the overarching goal of the course.

Data analysis is approached from the perspective of understanding statistics and their relationship to research rather than focus on mathematics, so underlying theory of statistics will be presented through content related to research and real case studies to facilitate learning.

The course introduces many of the important new ideas in data mining and machine learning, explains them in a statistical framework, and describes some of their applications to science and technology.

Real case studies will be presented and discussed, providing to student the opportunity to convey information about the statistical method and the results, especially the meaning of the results and how they relate to the question addressed.

During the course top scientists and field experts should be invited to give a thematic lesson on case studies or specific arguments.

Finally, to integrate the theory, a slot will be dedicated to provide the basic knowledge of R platform

COURSE SYLLABUS

Quick refresh of basic concepts [1 hour]

- Comparing Populations or Treatments. How to compute power of a test and choosing the sample size for testing population mean; How to compare the mean of two populations for independent samples, how to compute and interpret P value, How to use contingency table and the Chi-square test. [3 hours]

- General linear model: the analysis of variances and its extensions [2 hours]

- Practical on R statistics platform [2 hours] - (Dott. Mario Barbato)

LEARNING OBJECTIVES

A student who has met the objectives of the course is expected to:

- Choose the statistical method functional to the type of question to be addressed;
- Use statistical methods to extract meaning from large datasets
- Implement the method using software (e.g. R);
- Interpret the results and communicate the purposes of the analyses, the findings and their implications and in technically sound way;

ASSESSMENT PLAN

Team Project: a data set and will be provided to a team composed by two to three students and asked to employ techniques learned throughout this course to analyze the data set, interpret and report results

Final Exam: Written and oral examinations. Written examination consists of exercises designed to test the knowledge acquired during the course. Oral examination aims to evaluate the reasoning and arguing ability as well as to understand and interpret the statistic outputs.