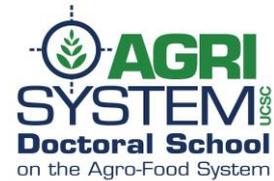




UNIVERSITÀ
CATTOLICA
del Sacro Cuore



APPLIED STATISTICS

Prof. Riccardo Negrini
Faculty of Agriculture, Food and Environmental Sciences

Course Objectives

This is a data-driven, applied statistics introductory course.

The overall course goal is to give the participants an introduction to statistical methods and data analysis, providing students with a conceptual and practical understanding of the application of statistics methods to agri-food data.

The course introduces the basic concepts of statistics, demonstrates methods of exploring, organizing, and presenting data, and introduces the fundamentals of probability. The course presents the foundations of statistical inference, including the concepts of parameters, confidence intervals, and hypothesis tests. Topics include an introduction to experimental design, linear regression, analysis of two-way tables, sample size and power calculation, permutation tests and resampling procedures.

All topics will be illustrated through examples and case studies drawn from broad areas of application along with agri-food production chains, and on real data from research studies, public database, on-field surveys and other sources. Upon course completion, students will appreciate and understand the role of statistics in agri food production sectors, develop a practical ability to apply appropriate statistical methods to analyze and summarize agri-food data, be able to interpret and report the results in appropriate table and graph.

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

Course Syllabus

The Role of Statistics and Data Analysis in Agri-food production

- Numerical Methods for Describing Data. How to describe data by summary statistics: measures of central tendency and variability
- Graphical Methods for Describing Data. How to describe agri-food data graphically for categorical data (pie chart, bar chart) and graphs for quantitative variables (histogram, stem-and-leaf plot, time plot, etc).
- Introduction to Probability and Distributions How binomial distributions and normal distributions are involved in statistics.
- Simple random sampling and sampling distribution of sample mean central limit theorem, normal approximation to the binomial
- Hypothesis Testing. How to set up Null and Alternative hypotheses, understanding Type

I and Type II errors, performing a statistical test for the population mean.

- An Introduction to experimental design. Steps for Planning, Conducting and Analyzing an Experiment
- 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
- Simple Linear Regression and Correlation. Understanding concepts related to linear regression models including, least squares method, correlation, Spearman's rank order correlation, inferences about the parameters in the linear regression model.
- Introduction to the analysis of variance and to the analysis of contingency tables. Understanding the F-test to compare the variances of two populations

Learning objectives

A student who has met the objectives of the course will be able to:

- Estimate and interpret simple summary statistics, (e.g. mean, standard deviation, variance, median and quartiles etc.)
- Apply appropriate simple graphical techniques (e.g. histograms, normal-score plots, box plots and scatterplots)
- Apply and interpret key statistical concepts, such as the formulation of models, parameter estimation, construction of confidence intervals and hypothesis testing
- Choose and apply and interpret simple statistical methods as regression and one way ANOVA
- Perform simple calculations using the program R

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 examinations. Written examination consists of exercises designed to assess the knowledge acquired during the course.

Recommended Text

- For learning statistical concepts, a functional textbook is by David M. Dietz, Christopher D. Barr, and Mine Cetinkaya-Rundel (2015). OpenIntro Statistics, American Institute for Mathematics. It is also available online for free as a pdf. (<https://www.openintro.org/stat/textbook.php>)
- For learning R, a usefull textbook is Nicholas J. Horton, Daniel T. Kaplan, and Randall Pruim (2015). A Student's Guide to R It is available online via this link: [https://cran.r-project.org/doc/contrib/Horton+Pruim+Kaplan_MOSAIC- StudentGuide.pdf](https://cran.r-project.org/doc/contrib/Horton+Pruim+Kaplan_MOSAIC-StudentGuide.pdf)

Websites of some interest:

- <http://onlinestatbook.com/rvls/index.html>
- <https://www.youtube.com/user/statisticsfun>