



---

# INFORMATICA E STATISTICA MEDICA

9 CFU - 1° Semester

## Teaching Staff

### ALFREDO PULVIRENTI

**Email:** apulvirenti@dmi.unict.it

**Office:** Stanza 35, Terzo Blocco Dipartimento di Matematica e Informatica.

**Phone:** 095-7383087

**Office Hours:** Martedì 10-11.

---

## LEARNING OBJECTIVES

### ▪ Bioinformatics

Objective of the course is the acquisition of methods for the analysis of biological sequences and structures and the capability of searching in biological databases (eg. Genes, sequences, functional domains). Starting from primary sequences of nucleic acids or proteins can hypothesize the function, evolutionary history and structure. The tools used to achieve these objectives are the public databases and the tools for the analysis and visualization of such kind of data.

**Knowledge and understanding:** Students will gain knowledge on methods for the analysis of biological sequences and for searching in biological databases. In particular students will be able to search on database of sequences and domains. Also, public databases available on NCBI will be presented together with software for the analysis and visualization of biological data. Finally, students will acquire the basic tools for the analysis of the transcriptome.

**Applying knowledge and understanding:** identify the appropriate tools to manipulate data and extract knowledge underlying; solve problems through the use of appropriate software in bioinformatics.

**Making judgments:** Through guided exercises, the students will acquire the basic skills necessary to deal with the analysis of new biological sequences, hypothesizing the function, study the transcriptome.

**Communication skills:** the student will acquire the necessary communication skills and expressive appropriateness in the use of technical language within the general framework of the analysis of biological data.

**Learning skills:** The course aims, as the goal, to provide students with the necessary basic theoretical methods and practices in order to address and solve problems concerning the analysis of biological data.

### ▪ Medical statistics

**General summary:**

The course aims to acquire the main basic concepts of probability and statistics.

General teaching training objectives in terms of learning outcomes:

**Knowledge and understanding:** The course aims to acquire skills to students about the description of statistical data; Understand the basic terms (population, sample, variable, etc.); Calculation and presentation of frequency distributions; data description with graphical methods; Calculation of central tendency and variability indices; Understand the basis of the assessment of probability of an event and of a random variable; Acquiring concepts related to inferential statistics such as estimates for interval confidence and hypothesis tests.

**Applying knowledge and understanding:** identify distributions appropriate to represent the knowledge underlying; solving problems of inferential statistics and probability.

**Making judgments :** Through concrete examples and case studies, the student will be able to independently develop solutions to specific problems and assess the suitability of a statistical inference problem and solution.

**Communication skills:** the student will acquire the necessary communication skills and expressive appropriateness in the use of technical language within the general framework of the analysis of data using statistical methods.

**Learning skills:** The course aims, as the goal, to provide students with the necessary theoretical and practical methods to address and solve problems independently in the statistical analysis of data.

## COURSE STRUCTURE

- **Bioinformatics**

Lectures

- **Medical statistics**

Lectures

---

## DETAILED COURSE CONTENT

- **Bioinformatics**

Introduction to relational Databases: • Data representation; • The Relational model; • Relational Algebra; • Database Design; • Access: Tables, queries, report, forms; • Examples and case studies. Biological and biomedical databases: - NCBI: The Entrez system. Methods for querying databases Gene, Nucleotide, Protein, PubMed, OMIM. - Special purposes databases: Protein Data Bank (PDB), Single Nucleotide Polymorphism DB (SNP), Gene Expression Omnibus (GEO), Genome, UCSC, ENSEMBL, Gene Ontologies DB, BioSystem Pathways, KEGG Patways, Patways Commons, Tarbase, Mirò, mirBase, Expressed Sequence Tag (EST). Locked

- **Medical statistics**

Introduction to probability and statistics: • Introduction to probability; • Events; • Definition of probability; • Conditioned events; • Bayes Theorem; • Discrete Random Variables; • Expectation, Variance, Covariance, Standard Deviation; • Bernoulli distribution; Binomial Distribution;

Hypergeometric Distribution; Negative Binomial Distribution; Geometric Distribution; Poisson Distribution; • Continuous Random Variables; • Uniform distribution; Exponential Distribution; Gaussian Distribution; • Examples and exercises; • Introduction to Descriptive Statistics; • The data, variables, variability, indices; • Statistical Inference: parameter estimation, statistical tests; • • Examples and exercises; • Introduction to R.

---

## TEXTBOOK INFORMATION

### ▪ **Bioinformatics**

- Anna Tramontano "Bioinformatica", Zanichelli
- Krane, Raymer. "Fondamenti di Bioinformatica" Pearson
- Jambeck, Gibas "Developing Bioinformatics Computer Skills, O'Reilly
- Pascarella-Paiardini "Bioinformatica" Zanichelli

### ▪ **Medical statistics**

Lantieri PB, Risso D, Ravera G: Statistica medica per le professioni sanitarie, II ed. McGraw-Hill

---