

Course Syllabus

General Information

Course Number: ININ 5405
Course Title: Statistical Methods in Bioinformatics
Credit-Hours: 3 hours

Course Description

Study and application of statistical methods related to the most important bioinformatics analyses including: sequence analysis, gene expression and phylogenetic trees. The methods under study include inferential statistics, statistical modeling, clustering analysis and Markovian processes.

Prerequisites

One of the following courses:

- InIn 4010. Probability and Statistics for Engineers
- Inin 5559. Engineering Statistics
- InCi 4136. Applied Statistics for Civil Engineering.
- Agro 5005. Biometrics
- EsMa 3101. Applied Statistics I.
- EsMa 4001. Mathematical Statistics I.
- EsMa 4006. Statistics for the Biological Sciences.
- Esta 3002. Business Statistics II.

Textbook

- **Mathur, S. (2010) Statistical Bioinformatics with R, s, New York City, New York: Elsevier (<https://www.elsevier.com/books/statistical-bioinformatics/mathur/978-0-12-375104-1>)**

References

- Nielsen, R. (2005) Statistical Methods in Molecular Evolution, New York City, New York: Springer Science and Business Media (<http://www.springer.com/us/book/9780387223339>)
- Draghici, S. (2012) Statistics and Data Analysis for Microarrays Using R and Bioconductor, Boca Raton, Florida :Chapman and Hall/CRC Press (<http://www.crcpress.com/product/isbn/9781439809754>)
- Durbin, R., Eddy, S., Krogh, A., Mitchison, G. (1998) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, New York City, New York: Cambridge University Press (<http://www.cambridge.org/us/academic/subjects/life-sciences/genomics-bioinformatics-and-systems-biology/biological-sequence-analysis-probabilistic-models-proteins-and-nucleic-acids>)
- Gentleman, R., Vincent, C., Hubert, W., Irizarry, R., Dudoit, S. (2005) Bioinformatics and Computational Biology Solutions Using R and Bioconductor, New York City, New York: Springer Science and Business Media (<http://www.springer.com/us/book/9780387251462>)
- Jones, N.C., Pevezner, P.A. (2004) An Introduction to Bioinformatics Algorithms, Cambridge, Massachusetts: The MIT Press (<http://mitpress.mit.edu/books/introduction-bioinformatics-algorithms>)

Course Syllabus

Purpose

This course has been designed for students in Engineering and the Sciences who are interested in the statistical and probabilistic aspects of bioinformatics. The main goal of the course is for the students to be able to use bioinformatics tools with a solid understanding of the statistical implications of their choices in exploration, analysis and conclusions. The course is an elective one.

Course Goals

After completing the course, the student should be able to:

Recognize and apply the appropriate statistical techniques that are used in the most common bioinformatics analysis procedures.

Interpret the results from statistically-based bioinformatics procedures and translate them into potential biological insight, to propose follow-up analyses or actions.

All students are expected to come to class on time, and prepared; do all assigned readings and related homework; actively participate in class discussions; and satisfy all assessment criteria to receive credit for the course.

Department and Campus Policies

Class attendance: Class attendance is compulsory. The University of Puerto Rico, Mayagüez Campus, reserves the right to deal at any time with individual cases of non-attendance. Professors are expected to record the absences of their students. Frequent absences affect the final grade, and may even result in total loss of credits. Arranging to make up work missed because of legitimate class absence is the responsibility of the student. (Bulletin of Information Undergraduate Studies).

Absence from examinations: Students are required to attend all required examinations. If a student is absent from an examination for a justifiable reason acceptable to the professor, he or she will be given a special examination. Otherwise, he or she will receive a grade of zero or "F" in the examination missed. (Bulletin of Information Undergraduate Studies)

Final examinations: Final written examinations must be given in all courses unless, in the judgment of the Dean, the nature of the subject makes it impracticable. Final examinations scheduled by arrangements must be given during the examination period prescribed in the Academic Calendar, including Saturdays. (See Bulletin of Information Undergraduate Studies).

Partial withdrawals: A student may withdraw from individual courses at any time during the term, but before the deadline established in the University Academic Calendar. (See Bulletin of Information Undergraduate Studies).

Complete withdrawals: A student may completely withdraw from the University of Puerto Rico, Mayagüez Campus, at any time up to the last day of classes. (See Bulletin of Information Undergraduate Studies).

Disabilities: After introducing and identifying himself/herself to the instructor and the institution as a student with disability, the student will receive reasonable accommodations in his/her courses and evaluations. For additional information, contact Services to Students with Disabilities at the office of the Dean of students using [787 – 265 – 3862](tel:787-265-3862) or [787 – 832 – 4040](tel:787-832-4040) exts. 3250, 3258).

Ethics: Any academic fraud is subject to the disciplinary sanctions described in Chapter VI of the revised General Student Bylaws of the University of Puerto Rico of the Board of Trustees. The professor will follow the norms established in articles 2.6 to 2.14 of the Bylaws.

Course Syllabus

General Topics

Lecture	Topic	Reading
1-5	Review of Probability and Statistics	Ch11
6-10	Introduction to Stochastic Processes	Ch 11
11	Evaluation of Basics	
12-17	Pairwise Alignment	Ch 2
18-23	Markov Chains and Hidden Models	Ch 3
24-29	Multiple Sequence Alignment	Ch 6
30	Second evaluation	
31-36	Gene Expression Data	Draghici Ch 2
37-42	Phylogenetic Trees	Ch 7
43-45	Presentation of Mini Projects	

Evaluation Strategies

	Quantity	Percent
Mini Projects	4	40-50%
Assignments	3	10-15%
Final Presentation	1	10-15%
Partial Exams	2	15-20%
Final Exam	1	5-10%
TOTAL:		100%