Optimization and Data Analysis in Biomedicine (syllabus)

Instructor	Contact Information
Andrey SHUVAEV Ph.D. in Biophysics, Associate professor, Institute of Biology and Biotechnology Siberian Federal University, Krasnoyarsk	Svobodny, 79 Room 32-13 Tel +7 391 206 2165 Email: AShuvaev@sfu-kras.ru, andrey.n.shuvaev@gmail.com

Course description

"Optimization and Data Analysis in Biomedicine" is the extensive course, which is designed to advance the student's ability to interpret experimental biomedical data and to construct primary mathematical descriptions of the phenomenon under investigation. It provides the wide comprehension of the statistical nature of biomedical data together with competence to operate the basic mathematical models in biology.

The most important part of this course is stating and verification of the statistically consistent hypotheses in biology. We will differentiate between several possible explanations of the given phenomenon from the experiment to find the most appropriate one.

The Optimization part of the course includes the inverse problem solution – determination of the kinetic parameters of the process with the experimental data and mathematical description available. Such an approach will allow to evaluate the characteristic velocities of the processes in the experiment.

The course is designed to cover a wide range of possible areas of biology and medicine. Grading this course will help the future specialist to construct and develop the explanation of the observed phenomena.

Course aims

- To help students to systematize and repeat their level of basic statistics, which is the important part for the major topics of the course
- To assist students to plan the experiment in the field of their interest
- To give students the understanding of how to analyze the hypotheses
- To introduce the basic principles of statistical and mathematical processing of the experimental data
- To give the introduction in optimization methods and it application in biological experiments

Course objectives

- To ensure that students are familiar with a basic statistical distributions
- To give students the optimal design methods for experiment planning
- To give students an appreciation of the testing algorithms and estimation criteria for hypothesis testing
- To provide students with the basic routines of statistical and mathematical model construction, through the data simulation and using the systems of ordinary differential equations

- To make students know and understand the principles of the parameters of the model estimation

Learning outcomes

By the end of the course, the students will be able to:

- select an optimal set of experimental conditions to obtain the maximum information
- compare the experimental outputs taken in different conditions
- construct the mathematical and/or statistical representation of the observed process from the experimental data
- determine the main characteristics of the experimental processes

Outline of content

Week	Lectures	Practice session / Assignments	Hours
1-4	Scientific method in biology and biomedecine. Experimental planning. How to collect the data	o Volume of the samples, repeatability o Models as the guess o Falsifiability of the theory, Theory of the partial success o Orthogonalization of the experimental conditions o Construction and operations with information matrix o Home assignment No 1	4/4 16
5-6	Basic statistics. Distribution of the collected outputs	o Idealized distributions o Experimental distributions o Statistical distributions o Home assignment No 2	2/2 16
7-8	Comparison of the samples. Hypotheses testing	o Testing algorithm o Test statistics o Estimation criterions o Home assignment No 3	2/2 16
9-12	Models types	o Correlation and dependence o Statistical models o Data simulation o Mathematical model o Home assignment No 4	3/3 16
13-15	Optimization and parameters stimation	o Meaning of optimization o Optimization techniques o Optimization quality and models comparison o Home assignment No 5	3/3 16
16		Final exam	2

Assessment and assessment method

Both midterm and final exams are in the form of a simulated experimental dataset. Problems for solution must be made with any suitable package: Matlab, Scilab, R etc.

Grade policy:

A (excellent work)	91-100 %
B (above average work)	81-90 %
C (average work)	71-80 %
D (below everage work)	50-70 %
F (failed work)	<50 %

Recommended reading

1. Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge University Press.

2. Montgomery, D. C. (2008). *Design and analysis of experiments*. John Wiley & Sons.

3. Manly, B. F. (2006). *Randomization, bootstrap and Monte Carlo methods in biology* (Vol. 70). CRC Press.