Statistics in Systems Biology and Programming in R

……..

Overview

Over the past 20 years the world of data analysis has been changed dramatically – we have entered the era of Big Data. Now amount of accumulated data significantly increases our abilities to analyse them. Sequencing the human genome, new DNA methylation chips providing the resolution of 900K probes per human genome, SNPs, different –omics data require the systematic development of new statistical approaches.

The aim of this course is to give systematic introduction to modern statistics and programming in R. The course will cover main topics from hypothesis testing to development of new and sophisticated algorithms for machine-learning. In all modules of this Introduction we will follow very practical approach oriented to immediate application of statistical techniques learned. The overarching goal of this course is to cover all main notions required to understand modern papers in biology and clinical analysis from statistical point of view.

Despite the short time we will discuss also widely used methods of graph analysis and approximate Bayesian calculations. Students will have an opportunity to apply the knowledge acquired with short algorithms programmed in R.

The Objectives of this course include:

1) Provide self-consisted introduction to modern techniques used in Systems Biology and analysis of clinical data.

2) Teach students how to use different statistical methods required to perform different statistical analysis, from simple hypothesis testing to more sophisticated classification or clustering tasks.

3) Teach student how to use learned algorithms using the programming language R.

The course will be divided into four modules that will be covered in a total of 16 lectures spanning over five working days in one week. The topics in Module A will expose the participants to the review of probability theory, conditional probability and Bayes theorem,

main distributions and Central Limit Theorem, T- and F- distributions. Module B is dedicated to the hypothesis testing. The topics in the module include P-value and basic tests, hypothesis testing with two and more samples, ANOVA testing, Kruskal-Wallis and Wilcoxon tests. Module C will cover data modelling. The topics include Correlation analysis, linear and logistic regression, Odds Ratio, clustering and graph analysis. In Module D, Approximate Bayesian Computations will be discussed. This includes Markov Chain Monte Carlo methods as well as a case study example: analysis of oncomarkers.

Modules	A: **Review of Probability Theory** B: **Hypothesis testing** C: **Data modelling** D: **Approximate Bayesian Computations** **Number of participants for the course will be limited to fifty.**
Schedule	**August 22 – 27, 2016,** 1 Credit course covering 16 lecture hours **Number of participants for the course will be limited to fifty.**
You Should Attend If…	▪ Undergraduate, Master or PhD level scholar who would like to be introduced to the new and growing interdisciplinary area of Network Science and Multi-Agent systems. ▪ Young and budding members of the faculty at various Engineering and Computer Science departments wanting to learn develop research programs in the respective departments. ▪ Scholars in governmental, industrial or consulting agencies interested in understanding the state of the art in this area.
Fees*	The participation fees for taking the course is as follows: **Participants from abroad : US $500** **Industry/ Research Organizations: Rs. 6000** **Academic Institutions: Rs. 1500 (For Students) and Rs. 3000 (For Post Docs and Faculties)** The above fee includes course materials, computer use for tutorials. The participants will be provided single-bed accommodation on payment basis *Please note that the following fees waiver will be provided if you register for our both the GIAN courses *i.e.* Biomathematics: from Gene Expression to Bone Mechanics (August 15-20, 2016) and Statistics in Systems Biology and Programming in R (August 22-27, 2016). **Participants from abroad : US $800** **Industry/ Research Organizations: Rs. 10,000** **Academic Institutions: Rs. 2500 (For Students) and Rs. 4500 (For Post Docs and Faculties)**

The Teaching Faculty

Dr. Oleg Blyuss holds a research position at the University College London in the frame of the CR-UK funded project devoted to the development of early diagnosis of ovarian cancer. He worked in the field of machine learning and its applications in biology and engineering. He developed modifications of fuzzy c-means algorithm for clustering based on nonlinear vector criterion. One of them was applied to the problem of the sandstone layer classification with respect to their burst hazard. Since 2012 Oleg was involved in the data analysis for identification of potential serum biomarkers of biliary tract cancer. Oleg performed the analysis of different ovarian cancer biomarkers and reported which of them have high prognostic possibilities. Together with the other collaborators on the project (Prof. Zaikin and Prof. Marino) Dr Blyuss developed and implemented a new biomarker longitudinal analysis algorithm and several methods, which have high practical importance for ovarian cancer detection.

Course Coordinator:

Dr. Sarika Jalan completed her PhD in Physics with specialization in nonlinear dynamics and Complex Systems from Physical Research Laboratory, India in 2005. She has six years post-doctoral experience at MPI-MiS, Leipzig , MPI-PKS and NUS, Singapore. During this period she worked on spectral properties of complex systems as well as applications to biological systems. Upon joining IIT Indore In December 2010, she established Complex Systems Lab, which focuses on fundamental research in the inter-disciplinary field of complex systems research, utilizing techniques from Physics, Mathematics, Bio-informatics and Computer Science. Using network theory, nonlinear dynamics and computational techniques, the lab on one hand works on developing tools pertaining to complex systems research and on other hand applies these techniques to real world systems coming from Biology and Social science.

Short Course Location

Indian Institute of Technology Indore

Duration

One week:
August 22 - August 27, 2016

Course Coordinator

Dr. Sarika Jalan
Associate Professor
Complex Systems Lab
Discipline of Physics
Indian Institute of Technology Indore
Simrol-453552, India

E-mail: sarikajalan9@gmail.com

....

http://www.iiti.ac.in/~sarika