Mel and Enid Zuckerman College of Public Health
University of Arizona

BIOS 648 Analysis of High Dimensional Data

Catalog Description: This course deals with the analysis of high dimensional data. It will cover multiple comparison, clustering and classification of high dimensional data, and regression methods involving high dimensional variables. Students will also learn the corresponding computer software. (3 units)

Course Topics:

- Introduction to R
- Multiple Comparison Methods
- Classification Trees
- Regression Trees
- Support Vector Machines
- Neural Networks

Course Objectives: During this course, students will:

- Learn modern methods for multiple comparisons, unsupervised clustering methods for high dimensional data, and classification and prediction models involving high dimensional information.
- Learn the corresponding data analysis computer software packages.
- Identify studies and data sets to which high dimensional methods should be applied.
- Identify appropriate methods to analyze high dimensional data.
- Use the statistical software package R to implement high dimensional methods.
- Interpret and critique medical and scientific journal articles which involve high dimensional data.

Learning Outcomes (Competencies Obtained): Upon completion of this course students will be able to:

1. Identify and apply appropriate statistical tools for high dimensional data, including inferential and predictive methodologies, for answering a particular research question
2. Recognize strengths and weaknesses of proposed statistical approaches
3. Communicate understanding of the assumptions necessary for a given statistical procedure as well as the ability to determine if the assumptions are met for a given study design or data set
4. Suggest preferred methodological alternatives to commonly used statistical methods when assumptions are not met
5. Manage data to handle a variety of practical problems in data format and structure
6. Apply appropriate statistical software in data analysis involving high dimensional outcomes
7. Demonstrate advanced competencies in areas of professional expertise and scholarship enabling advancement to further postgraduate study in statistics or biostatistics
8. Demonstrate the ability to identify, articulate and implement sound methodological and computational strategies for addressing scientific questions
9. Demonstrate the ability to communicate effectively in writing reports
10. Interpret and critique medical and scientific journal articles which involve high dimensional data