ST 810, Advanced Biostatistics, Fall 2014
Monday & Wednesday 3:00 - 4:15PM, SAS Hall 1108

Instructor: Dr. Wenbin Lu
Email: lu@stat.ncsu.edu
Homepage: http://www4.stat.ncsu.edu/~lu/ST810/ST810.html
Office: 5212 SAS Hall
Phone: (919) 515-1915
Office Hours: Wednesday 1:20 - 2:50PM (or by appointment)

Course Prerequisite: ST 521, 522.

Course Resources:

- References:

- Additional Online Materials: Please refer to course webpage:
  http://www4.stat.ncsu.edu/~lu/ST790/ST790.html.
  The contents will be updated regularly.

Course Description: Although theory of statistical inference has achieved a certain maturity in many fields of investigation, increasingly complex statistical models are emerging in biostatistics, economics, machine learning research, and in other areas. Over the past decades, developments in counting process and empirical process theories have proven to be powerful tools for working with these complex models and data structures. Semiparametric models, such as the proportional hazards model, that consist of both a parametric and a nonparametric component, are flexible and appealing because very few assumptions are made on the nonparametric part. However, this increased flexibility makes distribution theory quite challenging. Modern counting process and empirical process techniques are usually required. This is currently an extremely active as well as demanding area of statistical research.

This course will give a comprehensive review of counting process and empirical process theories, and their applications in various areas, including survival analysis and
semiparametric models. The topics include stochastic convergence in metric spaces, counting process-based martingale theory, Glivenko-Cantelli and Donsker theorems, semiparametric estimating equations, and nonparametric maximum likelihood estimation. In addition to examining numerous examples, the course will develop in each student the technical skills to enable application of the counting process and empirical process theories for studying semiparametric methods in statistics and biostatistics.

**Course Outline:**

1. Introduction and examples (1 week)
2. Counting process and martingale theory (4 weeks)
3. Cox's proportional hazards regression (2 weeks)
4. Other semiparametric regression models (1 weeks)
5. Empirical process and theories (6 weeks)
   - stochastic convergence in metric spaces
   - Glivenko-Cantelli and Donsker theorem
   - nonparametric maximum likelihood estimation
   - semiparametric efficiency
   - M-estimation and Z-estimation
6. Applications and case studies (2 weeks)

**Grading:** Letter grade will be given based on class participation (60%), homework (30%), and the final project (10%).

**Online evaluation:** Online class evaluations will be available for students to complete during the last two weeks of class. The website is https://classeval.ncsu.edu. Any comments and suggestion on this course are greatly appreciated.