PB HLTH C240F/STAT C245F
Statistical Genomics II

Sandrine Dudoit

Spring 2018

Syllabus

PB HLTH C240E–F/STAT C245E–F, Statistical Genomics I and II, both concern statistical methods and software for addressing inference problems that arise in genomic research.

Neither course is a prerequisite for the other.

Statistical Genomics II concerns high-throughput microarray and sequencing assays, for elucidating biological and medical questions relating to, for example: Transcription (RNA-Chip/RNA-Seq); protein-nucleic acid interactions, e.g., transcription factor binding sites (ChIP-Chip/ChIP-Seq); DNA methylation (methyl-Chip/methyl-Seq); DNA copy number (CGH DNA-Chip/DNA-Seq).

This semester, the course will first discuss the statistical analysis of meiosis and then focus on single-cell transcriptome sequencing (scRNA-Seq).

Practical Matters

- *Faculty instructor.*
  Sandrine Dudoit
  Website: `www.stat.berkeley.edu/~sandrine`
  E-mail: `sandrine@stat.berkeley.edu`
  Office hours: Tuesday, 14:00–15:00, 109 Haviland Hall

- *Reader.*
  Kelly Street
  Website: `statistics.berkeley.edu/people/kelly-street`
  E-mail: `kstreet@berkeley.edu`
  Office hours: TBA

- *Time and location.*
  Lecture: Tuesday and Thursday, 12:30–14:00, 334 Evans Hall
  Discussion: Wednesday, 11:00–12:00, 330 Evans Hall
  N.B. The discussion may be canceled this semester due to lack of graduate student instructor (GSI) support.

- *Registration information.*
  Public Health C240F, Class # 38326
  Statistics C245F, Class # 32959
  Units: 4

- *Grading policy.*
  50% assignments; 40% final project; 10% participation in lecture and discussion.

  Assignments involve both theory and data analysis using R and possibly other software.

  The final project consists of an abstract/proposal, written report, and poster or oral presentation on a topic that involves the application of statistical methods and software to address a particular biological or medical question.

- *References.*
  There is no required textbook. Lecture notes and references will be provided on the class website.

- *Prerequisites.*
  Statistics. STAT 201A–B (may be taken concurrently) or old version STAT 200A–B or consent of instructor.

  Computing. Some familiarity with the R language. Tutorials are available on the R Project website; references are posted on the class website.

  Biology. No formal training in biology is required; basic notions will be presented in class and references will be provided for further learning.

N.B. Please contact instructor if you do not satisfy the prerequisites. You are solely responsible for making up for any gaps in training.