Readings in Statistics and Probability: 
Analysis of Copy-Number Variation

University of Kentucky 
STA 715, Spring 2012 
Credit: 3.0

Meeting time: 1:00 p.m - 2:30 p.m., Wednesdays 
Room 205D, Multidisciplinary Sciences Building (MDS 205D)

Instructor: Patrick Breheny, Ph.D. 
Office: MDS 205D 
Phone: 218-2077 
e-mail: patrick.breheny@uky.edu 
Office hours: Whenever I’m in my office, or by appointment

Course description: We will be reading prominent articles pertaining to the analysis of copy-number variation data. In addition to reading articles specifically dealing with copy-number variation, we will also read some background material on some of the statistical methods involved. My goal is that at the end of the class, you will:

- Be familiar with the key articles and concepts in the analysis of copy-number variation data, as well as have a sense of the unanswered questions and current research directions
- Gain exposure to ideas and tools which can be applied to copy-number data, such as multiple testing corrections and false discovery rates, hidden Markov models, and the fused lasso
- Be more comfortable reading and discussing journal articles in statistics

Readings: I will provide .pdf copies of all the articles and textbook chapters we will be reading via Dropbox.

Prerequisite: STA 701, STA 703

Grading: Your grade will be based entirely on classroom participation and discussion. You should be fine as long as you actually read the articles and attempt to understand them. It is understandable and expected that you will still have some questions after reading the articles – this is the purpose of discussing them, after all. If I feel that your effort and participation are not at the “A” level, I will let you know.

Attendance: Obviously, since 100% of your grade is based on participation in discussion, attendance is very important. If you cannot attend one of our meetings, please let me know.
in advance. To make up the absence, you will be asked to play a larger role in the following week’s discussion. If you repeatedly miss class (more than 3 absences), it will be reflected in your grade.

**Electronic communication:** I will occasionally send e-mails to the class (to the account listed for you in the campus directory), so please check that account regularly.

**Complaints:** Students with suggestions or complaints should see me first, and if we cannot come to an agreement, I will direct you to the head of the department.

**Disabilities:** If anyone has a disability requiring special accommodations, please let me know as soon as possible, so that these arrangements can be made.

Copy-number variation is a phenomenon that has only recently been discovered, and there are many interesting aspects of its analysis that have not yet been explored. I hope that you will find both this specific topic interesting and also gain a sense of how new statistical methods are developed in response to the collection of new types of data and new inferential questions.

**Course outline:**

1. Fundamental aspects of copy-number variation

2. Landmark papers


3. Multiple comparisons and kernel methods


4. Circular binary segmentation


5. Hidden Markov Models


6. Fused lasso


7. Integrated analysis of SNP and CNV data
