Statistical Genomics

MAT 4376/5313 (MATH 6507)

Dr. David Bickel **Fall 2009** **M 11:30-13:00 FTX359 | W 11:30-13:00 FTX136**

Prerequisites: Either MAT 3172 & 3375 or instructor permission.

Contact: dbickel@uottawa.ca Office hours are by appointment (Room RGN 4510F).
Directions: Take the shuttle to 451 Smyth Rd. Follow the signs to the elevator, and take it to the fourth floor. Turn left and go all the way to the end of the hall, and enter the door to the OISB area.

Web page: http://tinyurl.com/kmet6w

1 Content

1.1 Course description

Statistical methods of analyzing high-dimensional data from transcriptional microarrays and from other genome-scale measurement technology. The class will cover both basic biological concepts and more advanced mathematical concepts needed to apply statistical decision and inference theory to genomics data.

The statistical methods will be explained using the same microarray examples throughout the class. Due to such repeated exposure, students should become familiar with microarray data, with the result that difficult statistical concepts can be understood without being obscured by unfamiliar biological concepts. The use of the same microarray data sets should also inculcate the awareness that the same data may be analyzed by more than one method.

Once the students become familiar with the statistical methods through applications to microarray data, other data types will be covered. Related data types include genome-wide association data as well as quantitative proteomics and lipidomics data.

1.2 Reading material

1.2.1 Required reading

- **Notes.** Lecture notes from the above web page. A draft version of the notes for a lecture will usually be available by 24 hours before the corresponding class begins, and exercises will be finalized within 24 hours after the class ends.
- **Textbook.** *An Introduction to Bayesian Analysis: Theory and Methods* by Jayanta K. Ghosh (Springer, 2006). The book is more Bayesian than the class: selected sections constitute only a small portion of the course's reading material, so consider sharing a copy of the textbook with classmates.
- **Supplement.** "A quick introduction to elements of biology - cells, molecules, genes, functional genomics, microarrays" by Alvis Brazma, Helen Parkinson, Thomas Schlitt, Mohammadreza Shojatalab. http://www.ebi.ac.uk/microarray/biology_intro.html

1.2.2 How to read the notes and textbook

Read actively, with pencil, paper, and eraser: http://home.sandiego.edu/~pmyers/textbook.html Helpful tips for reading math in general: http://www.math.ucdavis.edu/~tlewis/rfg/jclub/read_math.pdf

1.2.3 Introductions to measure theory

While the measure theory needed for the class is covered in the lecture notes, reading explanations by multiple authors tends to accelerate learning. The following resources are highly recommended.

- **Overview.** "A measure theory tutorial (Measure theory for dummies)" by Maya R. Gupta. Some of the terminology differs from that of the notes. For example, what Gupta calls *random variable* is what the instructor calls a *random quantity*. `https://www.ee.washington.edu/techsite/papers/documents/UWEETR-2006-0008.pdf`
- **Undergraduate-level book.** *Measure, Integral, and Probability* by M. Capinski and E. Kopp (Springer Undergraduate Mathematics Series, 2004).

2 Grading

2.1 Weights

2.1.1 Weights for the course grade

Component	Portion of course grade
Participation	5%
Pop quizes	10%
Midterm exam	25%
Final exam	60%

Participation means coming prepared to work out problems in front of the class for the benefit of all. Each quiz and exam is timed and may begin promptly at the scheduled start of a class.

Although students are not required to turn in homework, questions on the pop quizes may closely resemble homework problems and examples from the reading material, and the exams will also tend to favor those who have completed the assignments. Bringing a copy of the lecture notes to class is recommended in case of an "open note" quiz.

Extra credit worth a grade of 100% on one pop quiz will go to the first student who reports a mathematical error anywhere in latest version of the notes to dbickel@uottawa.ca. For example, if you are credited with discovering two mathematical errors and if your quiz grades at the end of class are 0%, 90%, and 80%, then your pop-quiz average would be $(2\,(100\%) + 0\% + 90\% + 80\%)\,/5$, hypothetically assuming that only three pop quizes were administered.

2.1.2 Weights for the final exam

Understanding assessed by the final exam (open notes)	Portion of exam
How to solve problems like those of midterm or of pre-midterm exercises & quizes	30% [[F]]
How to solve problems like those of post-midterm exercises & quizes	30% [[D]]
Subject matter of **main** post-midterm **objectives** *besides what is above*	25% [[B+]]
Content of lectures and assigned reading of lecture notes *besides what is above*	10% [[A]]
Content of assigned reading of textbook and supplement *besides what is above*	5% [[A+]]

Each letter in double brackets is the final-exam mark for a student who earns credit worth 90% of the portion of the exam at the corresponding row and above but no credit for any following rows.

2.2 University fraud policy

Even the *first* occurrence of plagiarism, cheating, or any other form of academic fraud will be reported to the dean of the Faculty in which the student is registered:
`http://www.commonlaw.uottawa.ca/index.php?option=com_content&task=view&id=3460&Itemid=46`

3 Assignments and exams

Date (2009)	Notes	Textbook	Supplement	Exercises
Sept. 9	N1	§§1.1, 1.2, Appendix D		t1.1, t1.2; n1
Sept. 14	N2	§1.3, p. 16 of §1.4.2, §1.4.3		t1.6, t1.7; n3
Sept. 16	N3		§1	n2, n4, n5
Sept. 21	N4	§§1.4.1, 1.5, 2.2		t1.11a, t1.13a,b; n6
Sept. 22		Last day for course changes.	N/A	N/A
Sept. 23	N5	§2.5		n7, n8, n9
Sept. 28	N6		§2.1	n10
Sept. 30	N7		§2.3	n11
Oct. 5	N8	§2.1	§2.4	
Oct. 7	N9	Review the above		Review the above
Oct. 12		**Thanksgiving Day**	N/A	N/A
Oct. 14	N10	§2.7.1		n12
Oct. 19	N11	§§3.1, 3.2		n13
Oct. 21		**Midterm exam**	N/A	N/A
Oct. 26	N12	§§9.6.0, 9.6.1		t2.17a,b
Oct. 28	N13	§2.12, 5.3, 9.1		
Nov. 2	N14			n14, n15
Nov. 4	N15			n16
Nov. 9	N16			n17, n18
Nov. 11	N17			n19
Nov. 16	N18	§§9.0, 9.2.0		n20
Nov. 18	N19	§§9.4, 9.5		
Nov. 23	N20			
Nov. 25	N21		§2.2	
Nov. 30	N22		§§3.1-3.4	n21
Dec. 2	N23		§4.3	n22
Dec. 7	N/A	Last day of class	§5.1	N/A
Dec. 9-23		**Final exam**	N/A	N/A

See Section 1.2.1 for access to the textbook (by Ghosh et al.), supplement (biology website), and notes (lecture notes). Exercises are from the textbook and specified in *chapter.exercise* format if preceded by "t" or from the lecture notes if preceded by "n."

A pop quiz administered on Monday may cover material related to assignments of any date prior to the date of the quiz, whereas one on Wednesday may cover material related to assignments of any date prior to Monday of the same week. This course information is tentative and expected to change twice every week.