BINF*6970 Statistical Bioinformatics

Winter 2023
Section(s): C01

College of Biological Science
Credit Weight: 0.50
Version 1.00 - January 06, 2023

---

1 Course Details

1.1 Calendar Description

This course presents a selection of advanced approaches for the statistical analysis of data that arise in bioinformatics, especially genomic data. A central theme to this course is the modelling of complex, often high-dimensional, data structures.

Restrictions: Restricted to Bioinformatics students.

1.2 Course Description

Welcome Message: Welcome to BINF*6970! I look forward to working with you this semester. Despite the conceptual nature of the course and all the mathematics I will strive to maintain a highly interactive structure of this course. Students in this course regularly "do things" rather than only listen. The best way to learn how to do statistical analyses is to do it!
Overview: The main goal of this course is to guide graduate students through the various statistical methods used in bioinformatics today. The course will start with some basics that students might still be familiar with from prior undergraduate and graduate courses. It will quickly move on to more complex approaches and statistical learning.
Curriculum Note: This course is complementary to others in the bioinformatics graduate program. The winter semester core bioinformatics courses are Genomic Methods (BINF*6110) and Biosequence Pattern Analysis (BINF*6420). In addition there are a number of complementary elective courses available (e.g., ANSC*6100 - Machine learning modelling)
Pre-Requisites: Students accepted into the Master of Bioinformatics and MSc in Bioinformatics programs should have the necessary background for this course. Programming experience in R is necessary for this course. Students are expected to have taken at least one course at the undergraduate level in statistics or biostatistics (or have the equivalent experience).
Course Format: This course will be held in person in a hybrid form. For the hybrid part we will be using using Zoom so that students can join, whether in class or remotely, regardless of their personal circumstances. Recordings of class will be made available but given the

interactive nature of the course just listening to those will not suffice. These are meant to repeat certain concepts on your own or if you miss a single class.

1.3 Timetable

Times: Tuesday and Thursday 10:00AM-11:20AM in SSC 1306

Start Date: January 10, 2023; End Date: April 6, 2023.

Delivery Mode: In-person with hybrid option

1.4 Final Exam

Exam time and location is subject to change. Please see WebAdvisor for the latest information.

---

2 Instructional Support

2.1 Instructor

Instructor: Dirk Steinke
Email: dsteinke@uoguelph.ca
Telephone: +1-519-824-4120 x53759
Office: CBG109
Office Hours: By appointment - virtually and in person

---

3 Learning Resources

3.1 Text

Lecture notes and recordings will be provided for recap and further study. Selected research articles and book chapters will also be shared via CourseLink as additional resources.

3.1 Software

All computational examples in the course (lectures, assignments) will be done in R and RStudio. Students are allowed to provide assignment solutions done with Python if they

chose to.

Note: Please bring your own laptop to every class for hands-on practice and software implementation of the methods learned in class.

3.1 CourseLink

All Course information and material (such as assignments, data sets, etc.) will be available on CourseLink. Students are responsible to check the website regularly for updated information and announcements. It is highly recommended to adjust announcement settings so that these are forwarded to email or text (or both). Any important changes (schedule etc.) are communicated through this system.

Students are expected to come prepared for class. Material will be posted on CourseLink in advance to allow for preparation.

---

4 Learning Outcomes

4.1 Course Learning Outcomes

By the end of this course, you should be able to:

1. understand statistical methods used in modern bioinformatics
2. analyze and visualize common types of bioinformatics data (mainly genomic information such as DNA sequences, microarray results, SNPs)
3. conduct reproducible analyses and use methods for statistical inference and statistical learning
4. understand and apply selected statistical concepts used in bioinformatics,
5. adapt the above skills to learn new methods and conduct new analyses

---

5 Teaching and Learning Activities

5.1 Lecture

Topics:	Weekly Lectures	Topics Covered	Not
	January 10	Probability theory	
	January 12	Regression & Correlation	**Qui**
	January 17	Probability distributions	
	January 19	Statistical inference	**Qui**
	January 24	Nonparametric Statistics	
	January 26	Bayesian Statistics	**Qui**
	January 31	Markov Chain Monte Carlo	
	February 2	Analysis of Variance	**Qui**
	February 7	Classification I	
	February 9	Classification II	**Assign** Due Frida
	February 14	Resampling methods: Cross validation	**Qui**
	February 16	Resampling methods: Bootstrap	
	Winter Break	No Classes Scheduled on February 21 and 23	**Qui**
	February 28	Linear models I	
	March 2	Linear models II	**Assignm** Due Friday

March 7	Beyond linearity	**Quiz**
March 9	Tree-based methods: Decision trees	
March 14	Tree-based methods: Random forest	**Quiz**
March 16	Support vector machines	
March 21	Deep learning	**Quiz**
March 23	Survival analysis	
March 28	Unsupervised learning: Ordination	**Assignm** Due Friday,
March 30	Unsupervised learning: Clustering	**Quiz**
April 4	Multiple testing	
April 6	Summary	**Assignm** Due Friday

---

6 Assessments

6.1 Marking Schemes & Distributions

Assignments 70%: 15% each for Assignments 1-3; 25% for Assignment-4.

Quizzes 30%: 3% each

6.2 Assignments

Assignments will consist of analyses of selected datasets. Students will turn in written reports on these, together with relevant graphics and conclusions. Programming code for each assignment is required and will be relegated to an appendix. The evaluation will be

based on the following criteria: amount and depth of the analysis, correct use of the statistical methods, correct and logical interpretations of the outcomes of the analyses, clarity and professional appearance of the text and graphics. The length of the text will not matter but brevity where possible is recommended. One assignment will be based on group activity.

6.3 Collaboration

While you are encouraged to discuss approaches to assignment questions with other students, the material turned in must be your own. Each individual assignment is intended to be solely the work of a single student (or the group) whose name appears on it.

6.4 Attendance

Although no explicit marks are given for class participation, attendance is crucial for successful completion of this course.

---

7 College of Biological Science Statements

7.1 Wellness

If you are struggling with personal or health issues:

- Counselling Services offers individualized appointments to help students work through personal struggles that may be impacting their academic performance.
- Student Health Services is located on campus and is available to provide medical attention.
- For support related to stress and anxiety, besides Health Services and Counselling Services, Kathy Somers runs training workshops and one-on-one sessions related to stress management and high performance situations.
  http://www.selfregulationskills.ca/

7.2 Personal information

Personal information is collected under the authority of the University of Guelph Act (1964), and in accordance with Ontario's Freedom of Information and Protection of Privacy Act (FIPPA) http://www.e-laws.gov.on.ca/index.html. This information is used by University officials in order to carry out their authorized academic and administrative responsibilities and also to establish a relationship for alumni and development purposes.

For more information regarding the Collection, Use and Disclosure of Personal Information policies please see the Undergraduate Calendar. (https://www.uoguelph.ca/registrar/calendars/undergraduate/current/intro/index.shtml)

7.3 Course Offering Information Disclaimer

Please note that course delivery format (face-to-face vs online) is subject to change up to the first-class day depending on requirements placed on the University and its employees by

public health bodies, and local, provincial and federal governments. Any changes to course format prior to the first class will be posted on WebAdvisor/Student Planning as they become available.

8 University Statements

8.1 Email Communication

As per university regulations, all students are required to check their e-mail account regularly: e-mail is the official route of communication between the University and its students.

8.2 When You Cannot Meet a Course Requirement

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons please advise the course instructor (or designated person, such as a teaching assistant) in writing, with your name, id#, and e-mail contact. The grounds for Academic Consideration are detailed in the Undergraduate and Graduate Calendars.

Undergraduate Calendar - Academic Consideration and Appeals
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml

Graduate Calendar - Grounds for Academic Consideration
https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml

Associate Diploma Calendar - Academic Consideration, Appeals and Petitions
https://www.uoguelph.ca/registrar/calendars/diploma/current/index.shtml

8.3 Drop Date

Students will have until the last day of classes to drop courses without academic penalty. The deadline to drop two-semester courses will be the last day of classes in the second semester. This applies to all students (undergraduate, graduate and diploma) except for Doctor of Veterinary Medicine and Associate Diploma in Veterinary Technology (conventional and alternative delivery) students. The regulations and procedures for course registration are available in their respective Academic Calendars.

Undergraduate Calendar - Dropping Courses
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml

Graduate Calendar - Registration Changes
https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/genreg-reg-regchg.shtml

Associate Diploma Calendar - Dropping Courses
https://www.uoguelph.ca/registrar/calendars/diploma/current/c08/c08-drop.shtml

8.4 Copies of Out-of-class Assignments

Keep paper and/or other reliable back-up copies of all out-of-class assignments: you may be asked to resubmit work at any time.

8.5 Accessibility

The University promotes the full participation of students who experience disabilities in their academic programs. To that end, the provision of academic accommodation is a shared responsibility between the University and the student.

When accommodations are needed, the student is required to first register with Student Accessibility Services (SAS). Documentation to substantiate the existence of a disability is required; however, interim accommodations may be possible while that process is underway.

Accommodations are available for both permanent and temporary disabilities. It should be noted that common illnesses such as a cold or the flu do not constitute a disability.

Use of the SAS Exam Centre requires students to make a booking at least 14 days in advance, and no later than November 1 (fall), March 1 (winter) or July 1 (summer). Similarly, new or changed accommodations for online quizzes, tests and exams must be approved at least a week ahead of time.

For Guelph students, information can be found on the SAS website
https://www.uoguelph.ca/sas

For Ridgetown students, information can be found on the Ridgetown SAS website
https://www.ridgetownc.com/services/accessibilityservices.cfm

8.6 Academic Integrity

The University of Guelph is committed to upholding the highest standards of academic integrity, and it is the responsibility of all members of the University community-faculty, staff, and students-to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have the responsibility of abiding by the University's policy on academic misconduct regardless of their location of study; faculty, staff, and students have the responsibility of supporting an environment that encourages academic integrity. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection.

Please note: Whether or not a student intended to commit academic misconduct is not relevant for a finding of guilt. Hurried or careless submission of assignments does not excuse students from responsibility for verifying the academic integrity of their work before submitting it. Students who are in any doubt as to whether an action on their part could be construed as an academic offence should consult with a faculty member or faculty advisor.

Undergraduate Calendar - Academic Misconduct
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml

Graduate Calendar - Academic Misconduct

https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml

8.7 Recording of Materials

Presentations that are made in relation to course work - including lectures - cannot be recorded or copied without the permission of the presenter, whether the instructor, a student, or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

8.8 Resources

The Academic Calendars are the source of information about the University of Guelph's procedures, policies, and regulations that apply to undergraduate, graduate, and diploma programs.

Academic Calendars
https://www.uoguelph.ca/academics/calendars

8.9 Disclaimer

Please note that the ongoing COVID-19 pandemic may necessitate a revision of the format of course offerings, changes in classroom protocols, and academic schedules. Any such changes will be announced via CourseLink and/or class email.

This includes on-campus scheduling during the semester, mid-terms and final examination schedules. All University-wide decisions will be posted on the COVID-19 website (https://news.uoguelph.ca/2019-novel-coronavirus-information/) and circulated by email.

8.10 Illness

Medical notes will not normally be required for singular instances of academic consideration, although students may be required to provide supporting documentation for multiple missed assessments or when involving a large part of a course (e.g.. final exam or major assignment).

8.11 Covid-19 Safety Protocols

For information on current safety protocols, follow these links:

- https://news.uoguelph.ca/return-to-campuses/how-u-of-g-is-preparing-for-your-safe-return/
- https://news.uoguelph.ca/return-to-campuses/spaces/#ClassroomSpaces

Please note, these guidelines may be updated as required in response to evolving University, Public Health or government directives.
