# STAT*4350/6821 :  Applied Multivariate Statistical Methods

# Course Outline

**Instructor:** Ayesha Ali          **Email:**  aali@uoguelph.ca
**Office**:      MACN 509          **Phone:** Ext. 53896

**Office Hours:**  Wed.        1:00 – 2:00pm  (subject to change)
**Lectures:**        Tu, Th    1:00 – 2:20pm   ROZH 107

**Pre-requisites:**  (1 of MATH*1160, MATH*2150 or MATH*2160), STAT*3110, STAT*3240

**Required Text:** *Applied Multivariate Statistics with R*, D. Zelterman, Springer, 2015. (online in library)

**Recommended Texts:**
*Applied Multivariate Statistical Analysis*, 5th Ed., Johnson and Wichern, Preston Hall, 2002.
*Methods of Multivariate Analysis*, 3rd Ed., Renchner and Christensen, Wiley, 2012. (online in library)

**Calendar Description:** This course introduces the multivariate normal, and Wishart and Hotelling's T-square distributions. Topics covered include: statistical inference on the mean vector, canonical correlation, multivariate analysis of variance and covariance, multivariate regression, principal components analysis, and factor analysis. Topics will be illustrated using examples from various disciplines.

**Course Objectives and Outcomes:**  The objective of this course is to acquaint students with the basic ideas, applicability, and methods of multivariate data analysis. Students who have successfully completed this course will be expected to:

- Have a general knowledge and understanding of many of the key concepts, theoretical approaches and assumptions needed for dealing with multivariate problems;
- Derive some fundamental classical results of multivariate analysis;
- Recognize types of problems (e.g. classification versus clustering) and the appropriate method(s) of analysis;
- Analyze multivariate data using statistical software (e.g. using R and related R packages);
- Develop independent and critical thinking skills with respect to multivariate data;
- Describe statistical methods and analysis results for multivariate data to a non-statistician in a written report with appropriate tables and figures.

**Course Description**: We will begin with a brief review of matrix algebra and random vectors. We will then move on to the multivariate normal distribution and making inferences about one or more means. Relevant distributions also include the Hotelling's T-squared distribution and the Wishart distribution. Statistical methods of analysis include multivariate analysis of variance and covariance (MANOVA, MANCOVA), multivariate regression, and dimension reduction methods including principal components analysis, factor analysis and canonical correlation analysis. Other important multivariate techniques are discrimination, classification and clustering methods. Time permitting, or through projects/presentations,

we will look at advanced topics such as multivariate adaptive regression splines, projection pursuit regression, graphical Markov models, and dealing with discrete multivariate data.

| Course Assessment: | STAT*4350 | STAT*6821 |
|---|---|---|
| Assignments (4, equally weighted) | 20% | 20% |
| Midterm (Tues. Oct. 23, in-class) | 20% | 10% |
| Project | 20% | 25% |
| Presentation (graduate students only) | – | 10% |
| Final Exam (Fri., Dec. 14, 8:30-10:30am) | 40% | 35% |

**Project:** The project will consist of a written statistical report of a particular application of a multivariate technique to analyse a data set. **You are responsible for proposing your own project topic and finding suitable data.** You will have to submit a project proposal to the instructor approximately one month before the project due date.

Usually, a project will require data exploration and a clear understanding of the materials presented in class or any reading assignments. The project is an extremely important tool, not only to help you understand the course subject matter, but also to help you develop a critical view of data analysis and refine your written communication skills. Your report must document the procedures used in your project and relevant findings. **Raw computer output is not acceptable**. Further details for the project will be given out later in the semester.

**Presentation:** (Graduate students only). The presentation will involve finding a recent article or book chapter that covers an advanced topic in multivariate analysis. You will have to read the article/chapter, understand it, and present it to the class in a twenty to twenty-five minute talk (with possibly more time if the paper requires it). The dates and order of presentations will be determined later on in the semester. BEFORE you start working on the paper, inform the instructor of your choice. The instructor may deem the article not suitable for this presentation and may suggest an alternate article. However, **you are encouraged to select the paper yourself.**

**Assignments:** Assignments will be posted on Courselink and due in class approximately every other week. Students are encouraged to discuss assignment questions with each other, but you are expected to complete the assignments yourself and to submit your own work. See below for the University of Guelph policies on Academic Misconduct.

**Schedule of Important Dates:**

| | |
|---|---|
| Th. Sept. 27 | Assignment 1 due |
| **Tu. Oct. 9** | **NO CLASS** – Fall Study Break |
| Th. Oct. 11 | Assignment 2 due |
| **Tu. Oct. 23** | **MIDTERM EXAM**, in-class |
| Fri. Oct. 26 | Project proposal |
| Th. Nov. 1 | Assignment 3 due |
| Th. Nov. 15 | Assignment 4 due |
| **Th. Nov. 29** | **Final Project due for undergraduates (STAT*4350)** and last day of lecture |
| Tu. Dec. 4 | Final project due for graduates (STAT*6821) |
| **Fri. Dec. 14** | **FINAL EXAM   8:30 – 10:30**, room TBA |

**Late Policy:** The instructor reserves the right to not accept late assignments, projects or presentations. If you cannot submit such assessments on time, you are advised to notify the instructor as soon as possible to discuss the matter. In some instances, a late penalty may apply.

**Computing:** We will make use of computing since data analysis is a central part of the course material. It is assumed that you already have some experience with the R statistical language, though not necessarily with respect to the topics covered in this course. However, there are several online tutorials available on the CRAN website to help you in this regard (http://cran.r-project.org/). The most recent version of R is also freely available from this site. One of the recommended textbooks is also tailored to multivariate analysis using R.

**Courselink:** Class announcements, assignments, handouts and partial lecture notes will be posted on Courselink and students are expected to regularly check the course website on Courselink.

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

**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. See the undergraduate calendar for information on regulations and procedures for Academic Consideration.

**Drop Date**: The last day to drop this course is Friday, November 2, 2018. The regulations and procedures for Dropping Courses are available in the Undergraduate Calendar.

**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.

**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 book their exams at least 7 days in advance, and not later than the 40th Class Day.

More information: www.uoguelph.ca/sas

**Academic Misconduct**: 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 discourages misconduct. 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.

[The Academic Misconduct Policy is detailed in the Undergraduate Calendar](#).

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

**Resources:** The [Academic Calendars](#) are the source of information about the University of Guelph's procedures, policies and regulations which apply to undergraduate, graduate and diploma programs.