



CUMMING SCHOOL OF MEDICINE
GRADUATE COURSE OUTLINE

COURSE TITLE: Advanced Statistical Modeling			
Course	DATA 621		
Pre/Co-Requisites	Data Science 602, 603 and admission to the Post-baccalaureate Certificate in Fundamental Data Science and Analytics or the Post-baccalaureate Diploma in Data Science and Analytics		
Faculty	Cumming School of Medicine, Graduate Science Education		
Instructor Name(s)	Alberto Nettel-Aguirre	Email	anettela@ucalgary.ca
Office Location	C4-435 Alberta Children's Hospital	Office Hours	By Appointment
Instructor Email Policy	Only emails coming from ucalgary addresses will be replied to. Questions that can be answered by reviewing notes or recommended readings will take really low priority to be replied to.		
Telephone No.	403 955 7880		
TA Name, if applicable	Frank Lee	Email	seungwon.lee@ucalgary.ca
Class Term, Days	Winter 2020 Tuesdays (there will be one class on Monday as coordinated with DATA 623)		
Class Times	17:00-19:50		
Class Location	Centre for Health Informatics, TRW 5E16 (Gosling Room)		

COURSE INFORMATION/DESCRIPTION OF THE COURSE
A more in-depth jump into the fundamental statistical methods used in health data science including interpretation and communicating the results of these methods. Exploration of modelling using an epidemiological paradigm including the assessment for modification and confounding. Introduces fundamental health research methods including study design and its effect on the modeling to be used. Topics include: Review of Multiple Linear Regression with Model Diagnostics and Model selection, Modeling binary outcomes: Contingency tables, Odds, Odds Ratios; modeling with more variables through Logistic Regression. Count data and rates, incidence rates, rate ratios and modeling through Poisson regression and Negative Binomial Regression. Modeling in the non-independent/clustered data setting: GEE, Mixed effects Models. Catching and modeling non-linear relationships in a linear model setting (Categorization, Smoothing Splines, etc.)
LEARNING RESOURCES/REQUIRED READING
You will be using R Studio in this course. Please see the download instructions for R Studio. Select the <i>R Studio Desktop</i> option with the <i>Open Source License</i> (First column option on the page).

No textbook is required per se. Readings and textbooks are recommended resources unless instructed in lecture or specifically.

Recommended resources:

1. Rothman, K., Greenland, S., & Lash, TL. (2008). *Modern Epidemiology, 3rd Edition*. Philadelphia, PA: Lippincott Williams & Wilkins
2. Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2005). *Statistics for biology and health. Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. New York, NY, US: Springer Publishing Co.
3. Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text, Third Edition*. New York, NY: Springer Science+Business Media, LLC.
4. McCullagh, P., Nelder, J. (1989). *Generalized Linear Models, Second Edition*. Chapman & Hall.
5. Gareth James • Daniela Witten • Trevor Hastie Robert Tibshirani, *An Introduction to Statistical Learning with Applications in R*: Springer New York Heidelberg Dordrecht London.
6. Wickham and Golemund, *R for Data Science*: O'Reilly Media.
7. Ben Baumer, Daniel T. Kaplan, and Nicholas J. Horton, *Modern Data Science with R*: CRC Press LLC, 2016
8. Venables, W., Ripley, B. (2003). *Modern Applied Statistics with S*. Springer.

COURSE OBJECTIVES/LEARNING OUTCOMES

The objectives of this course are to have students be able to understand the impact and nuances of statistical modeling as it corresponds to analyzing and interpreting results for Health data. Students should be able to know the importance and concepts of confounding and effect modification as well as techniques to assess for each, the issues these represent at time of interpretations of effects and the roles they play in modeling. Students will learn about techniques for modeling categorical outcome data, count data, time to event data, as well as learn about techniques used in case of non independent data.

CUT POINTS FOR GRADES

This course adheres to the grading system outlined in the University of Calgary, Faculty of Graduate Studies Calendar. Grades of A+ and A are not distinguished in the calculation of GPAs. Percentage/letter grade conversion used for this course is as follows

Grade	Grade Point Value	Percentage Conversion	Graduate Description
A+	4.00	95-100	Outstanding
A	4.00	90-94	Excellent – superior performance showing comprehensive understanding of the subject matter

A-	3.70	85-89	Very Good Performance
B+	3.30	77-84	Good Performance
B	3.00	72-76	Satisfactory Performance
B-	2.70	68-71	Minimum Pass for Students in the Faculty of Graduate Studies
C+	2.30	63-67	All grades below 'B-' are indicative of failure at the graduate level and cannot be counted toward Faculty of Graduate Studies course requirements

Assessment Components: The University policy on grading related matters is outlined in the 2019-2020 Calendar .			
Assessment Methods	Description	Weight %	Due Date <u>and</u> Time
Assignments (8)	Assignments will be aimed at testing understanding of course material (see below). Part of the assignment and learning is to be able to apply techniques and interpretations in different settings and hence not just to repeat exactly what was seen in class. The aim of assignments is to broaden the students' comprehension and application of material.	60	At beginning of class on the day scheduled
In class participation/presentation	There will often be a hands-on section of the class, participating on this and presenting results from such is expected.	10	
Project and presentation(1)	The project will aim at giving the student the opportunity to showcase the understanding of data and how it relates to situation at hand, explanation and reasoning for approach to analysis, interpretation and limitations of results and conclusions in context. The project will be assigned during the term. The project shall be completed by April 22, 2020. The purpose of the project is for students to be able to fully apply skills acquired during the course, namely interpretation of a research question and its translation into a statistical problem, the appropriate modeling and analysis to	20 project handout 10 presentation	Tentative date April 29, 2020

	be done, actual performing of the analysis, with all assumption checks and interpretation, as well as reporting of results. Whether it is based on your own data set or on data obtained from other sources. Each student (maybe in a group) must make a full intellectual contribution to the project. The project will consist of several components 1) A typed write-up containing: introduction, methods used to complete the project, results of the analysis, discussion detailing the conclusions of the analysis in R markdown. 2) The computer R code with complete detailed description of the codes. and the resulting output from the computer R code. The project will be presented in a short oral presentation.		
--	---	--	--

ASSESSMENT AND EVALUATION INFORMATION

ATTENDANCE AND PARTICIPATION EXPECTATIONS:

See [section E.3](#) regarding attendance. Apart from what is explained here, students are responsible for the material indicated, delivered and suggested during class. Not attending a class is not an excuse for not covering the material yourself. Any in class evaluation is still governed by the E.3 and M.1 sections.

GUIDELINES FOR SUBMITTING ASSIGNMENTS:

You will have to respond to 5 assignments which will be based on the material covered in class and in indicated readings. Quizzes will be created under the assumption that such assignments have been completed in full and that you have a complete comprehension of the course material, as such, you are fully responsible for reviewing your assignments and making sure the material is understood .

Your assignment must be your own work and must be completed in **R Markdown/Jupyter Book** and submitted as either an .html or .pdf file; in the case of specific coding requirements the .rmd if using R markdown. The assignments shall be submitted via D2L.

FINAL EXAMINATIONS:

There is no final examination per se, the project will be the final evaluation of the course.

EXPECTATIONS FOR WRITING:

The expectation is that all written and presented content be your own. Refer to [section K](#) of the calendar for issues with academic misconduct such as plagiarism and others.

LATE AND/OR MISSING ASSIGNMENTS:

In the event that a student misses the midterm or any course work due to illness, supporting documentation, such as a medical note or a statutory declaration will be required (see [Section M.1](#); for more information regarding the use of statutory declaration/medical notes, see [FAQ](#)). Absences must be reported within 48 hrs. The regulations of the Faculty of Science pertaining to this matter are found in the Faculty of Science area of the Calendar in [Section 3.6](#). It is the student's responsibility to familiarize themselves with these regulations. See also [Section E.3](#) of the University Calendar.

Any other late or missing work not covered by the above sections will not be accepted and hence, **there will be no makeup assignments/quizzes/projects.**

Is a passing grade on a particular component essential to pass the course as a whole? No

		COURSE TIMETABLE	
Course Schedule Date	Topic & Reading	Instructor	Assignments/Due Dates & Times
Jan 14, 2020	Course overview discussion. Health research methods: Hierarchy of evidence, Stating a Research question/project, Study design, Biases, confounding, effect modification	Nettel-Aguirre	Assignment #1, Due Jan 21 17:00
Jan 20, 2020	Review of linear regression in the context of previous epi concepts.	Nettel-Aguirre in lieu of Mar 10 or 24	No Assignment
Jan 21, 2020	Review of Logistic regression in the context of previous epi concepts.	Nettel-Aguirre	Assignment #2, Due Jan 28 17:00
Jan 28, 2020	Regression methods for polytomous outcome variables.	Nettel-Aguirre	Assignment #3, Due Feb 4 17:00
Feb 4, 2020	Count Data: Basic concepts, Poisson distribution, rates, rate comparisons	Nettel-Aguirre	Assignment #4, Due Feb 11 17:00
Feb 11, 2020	Regression Models for Count data	Nettel-Aguirre	Assignment #5, Due Mar 3 17:00
Feb 18, 2020	No class, Reading week	Nettel-Aguirre	No assignment
Feb 25, 2020	Regression Models for count data (cont.)	Nettel-Aguirre	Assignment #5, Due Mar 3 17:00
Mar 3, 2020	Time to event data, concepts, survival, hazard, Kaplan Meier	Nettel-Aguirre	Assignment #6 Due Mar 10 17:00
Mar 10 or 24, 2020	Cox Regression Model for time to event data	Guest Lecturer	Assignment #7 Due Mar 31 17:00
Mar 31, 2020	Non-independent observations (paired data, matched data, clustered data, longitudinal data). Modeling of such data.	Nettel-Aguirre	
Apr 7, 2020	Modelling of non-independent data	Nettel-Aguirre	Assignment #8 Due April 14 17:00
Apr 14, 2020	Other ways to address non-linear relationships in regression models.	Nettel-Aguirre	

INTERNET AND ELECTRONIC COMMUNICATION DEVICE INFORMATION



Cell phones must be turned off in class unless otherwise arranged with the instructor.

The use of laptop and mobile devices is acceptable when used in a manner appropriate to the course and classroom activities. Students are to refrain from accessing websites that may be distracting for fellow learners (e.g. personal emails, Facebook, YouTube). Students are responsible for being aware of the University's Internet and email use policy, which can be found at <https://www.ucalgary.ca/policies/files/policies/electronic-communicationspolicy.pdf>.

MEDIA AND RECORDING IN LEARNING ENVIRONMENTS

Media recording for lesson capture

The instructor may use media recordings to capture the delivery of a lecture. These recordings are intended to be used for lecture capture only and will not be used for any other purpose. Although the recording device will be fixed on the Instructor, in the event that incidental student participation is recorded, the instructor will ensure that any identifiable content (video or audio) is masked, or will seek consent to include the identifiable student content to making the content available on University approved platforms.

Media recording for assessment of student learning

The instructor may use media recordings as part of the assessment of students. This may include but is not limited to classroom discussions, presentations, clinical practice, or skills testing that occur during the course. These recordings will be used for student assessment purposes only and will not be shared or used for any other purpose.

Media recording for self-assessment of teaching practices

The instructor may use media recordings as a tool for self-assessment of their teaching practices. Although the recording device will be fixed on the instructor, it is possible that student participation in the course may be inadvertently captured. These recordings will be used for instructor self-assessment only and will not be used for any other purpose.

Student Recording of Lectures

Audio or video recording of lectures is prohibited except where explicit permission has been received from the instructor.

UNIVERSITY OF CALGARY POLICIES AND SUPPORTS

ACADEMIC ACCOMMODATIONS

Students seeking an accommodation based on disability or medical concerns should contact Student Accessibility Services; SAS will process the request and issue letters of accommodation to instructors. For additional information on support services and accommodations for students with disabilities, visit www.ucalgary.ca/access/. Students who require an accommodation in relation to their coursework based on



a protected ground other than disability should communicate this need in writing to their Instructor. The full policy on Student Accommodations is available at <http://www.ucalgary.ca/policies/files/policies/student-accommodation-policy.pdf>

IMPORTANT INFORMATION

Any research in which students are invited to participate will be explained in class and approved by the appropriate University Research Ethics Board

COPYRIGHT

It is the responsibility of students and professors to ensure that materials they post or distribute to others comply with the Copyright Act and the University's Fair Dealing Guidance for Students (library.ucalgary.ca/files/library/guidance_for_students.pdf). Further information for students is available on the Copyright Office web page (<https://library.ucalgary.ca/copyright>)

A NOTE REGARDING INSTRUCTOR INTELLECTUAL PROPERTY

Generally speaking, course materials created by professor(s) (including course outlines, presentations and posted notes, labs, case studies, assignments and exams) remain the intellectual property of the professor (s). These materials may NOT be reproduced, redistributed or copied without the explicit consent of the professor. The posting of course materials to third party websites such as note-sharing sites without permission is prohibited. Sharing of extracts of these course materials with other students enrolled in the course *at the same time* may be allowed under fair dealing

ACADEMIC INTEGRITY

The Cumming School of Medicine expects intellectual honesty from its students. Course participants should be aware of University policies relating to Principles of Conduct, Plagiarism and Academic Integrity. These are found in the printed Faculty of Graduate Studies Calendar, or online under Academic Regulations in the Faculty of Graduate Studies Calendar, available at [Faculty of Graduate Studies Academic Regulations](#)

ACADEMIC MISCONDUCT

For information on academic misconduct and its consequences, please see the University of Calgary Calendar at <http://www.ucalgary.ca/pubs/calendar/current/k.html>

EMERGENCY EVACUATION AND ASSEMBLY POINTS

Assembly points for emergencies have been identified across campus. The primary assembly points for South Campus (Health Science Centre (HSC); Health & Research Innovation Centre (HRIC); Heritage Medical Research Building (HMRB) and Teaching, Research and Wellness (TRW)) are:

- HSC and HMRB: HRIC Atrium (alternate assembly point is Parking Lot 6)
- HRIC: HMRB Atrium (alternate assembly point is Parking Lot 6)
- TRW: McCaig Tower (alternate assembly point is HMRB – Atrium)

APPEALS

If there is a concern with the course, academic matter or a grade, first communicate with the instructor. If these concerns cannot be resolved, students can proceed with an academic appeal, as per Section N of the Faculty of Graduate Studies Calendar. Students must follow the official process and should contact the



Student Ombuds Office (<http://www.ucalgary.ca/provost/students/ombuds>) for assistance with this and with any other academic concerns, including academic and non-academic misconduct

THE FREEDOM OF INFORMATION AND PROTECTION OF PRIVACY (FOIP) ACT

This course is conducted in accordance with the Freedom of Information and Protection of Privacy Act (FOIP) and students should identify themselves on written assignments (exams and term work.) by their name and ID number on the front page and ID on each subsequent page. Assignments given by you to your course instructor will remain confidential unless otherwise stated before submission. The assignment cannot be returned to anyone else without your expressed permission to the instructor. Grades will be made available on an individual basis and students will not have access to other students' grades without expressed consent. Similarly, any information about yourself that you share with your course instructor will not be given to anyone else without your permission

WELLNESS AND MENTAL HEALTH RESOURCES

The University of Calgary recognizes the pivotal role that student mental health plays in physical health, social connectedness and academic success, and aspires to create a caring and supportive campus community where individuals can freely talk about mental health and receive supports when needed. We encourage you to explore the excellent mental health resources available throughout the university community, such as counselling, self-help resources, peer support or skills-building available through the SU Wellness Centre (Room 370, MacEwan Student Centre), <https://www.ucalgary.ca/wellnesscentre/services/mental-health-services> and the Campus Mental Health Strategy website <https://www.ucalgary.ca/mentalhealth/>

SUPPORTS FOR STUDENT LEARNING, SUCCESS, AND SAFETY

Student Ombudsman: The Student Ombuds' Office supports and provides a safe, neutral space for students. For more information, please visit www.ucalgary.ca/ombuds/ or email ombuds@ucalgary.ca

Student Union: The SU Vice-President Academic can be reached at (403) 220-3911 or suvpaca@ucalgary.ca; Information about the SU, including elected Faculty Representatives can be found here: <https://www.su.ucalgary.ca>

Graduate Student's Association: The GSA Vice-President Academic can be reached at (403) 220- 5997 or gsa.vpa@ucalgary.ca; Information about the GSA can be found here: <https://gsa.ucalgary.ca>

SAFEWALK

Campus security will escort individuals, day or night, anywhere on campus (including McMahon Stadium, Health Sciences Centre, Student Family Housing, the Alberta Children's Hospital and the University LRT station). Call 403-220-5333 or visit <http://www.ucalgary.ca/security/safewalk>. Use any campus phone, emergency phone or the yellow phone located at most parking lot pay booths. Please ensure your personal safety by taking advantage of this service.