Land Acknowledgement

We respectfully acknowledge the Syilx Okanagan Nation and their peoples, in whose traditional, ancestral, unceded territory UBC Okanagan is situated.

DATA 410: Regression and Generalized Linear Models

Faculty: Irving K. Barber Faculty of Science
Department: Computer Science, Mathematics, Physics, and Statistics
Instructor(s): John R.J. Thompson
Instructor(s) Email: john.thompson@ubc.ca
Duration: Term 2 Winter 2022
Delivery Modality: In-Person
Course Location: ART 218
Course Days: Mon/Wed
Class Hours: 8:00 AM - 9:30 AM
Office hours: Mondays 9:30-11:00 am, ASC 348

Other Instructional Staff

1	Tutorial Instructor and Teaching Assistant, Nima Eslami, nima.eslami@ubc.ca
2	Laboratory time: 3:00pm - 5:00pm Thursdays online
3	Zoom URL: TBA

Course Description

Regression, linear models, generalized linear models, additive models, generalized additive models, mixed models, theory and numerical performance.

Pre-reqs: All of MATH 221, STAT 303, DATA 311.

Course Format

The course will be delivered via in-class lectures complemented by out-of-class readings, and programming and writing assignments. Pre-lecture readings will be given before the class, and students are expected to read them first before attending the class. Students with only a cursory knowledge of R, Markdown, and/or probability should study the textbook appendices carefully. Course content will be posted online via Canvas. Midterm break and other calendar dates can be found at http://okanagan.students.ubc.ca/calendar/. There will also be a graded project and in-class/in-person exams (see schedule below).

Course Overview, Content and Objectives

Regression analysis is the study of determining the characteristics of a relationship between an outcome variable and sets of predictor variables. The methodologies taught in this course are widely used in statistical and data science applications, and in many fields of science. The course begins with a thorough and rigorous review of multiple regression, including parameter estimation, testing, influence and multicollinearity diagnostics, modelling with factors. The remainder of the course will serve as an introduction to generalized linear models including the most commonly used models and modern extensions through spline smoothing and generalized additive models.

Learning Outcomes

Upon successful completion of this course, students will be able to...

conduct parameter estimation, statistical inference, residual analysis, model diagnostics and variable selection for generalized linear, nonlinear and generalized additive regression models. Students will not only be able to carry out regression analyses on a variety of data types, and but they will have a sufficient grasp of the underlying mathematics so that extensions to more sophisticated methodologies will be accessible without need for further instruction. Such a student should be able to consult the modern statistical literature and, with some effort, learn how to carry out analyses of types not covered in the course, such as quantile regression, survival analysis, time series analysis and hierarchical linear modelling. Students will learn to use the statistical language R to conduct regression analysis on data to answer scientific questions and to create dynamic documents with embedded R-code to foster reproducible research through their assignments and projects.

Assessments of Learning

• Assignments (6), 10%, Due on: January 23rd, February 6th, February 27th, March 13th, March 27th, April 13th.
• Project, 35%, Proposal due March 6th, Report due April 13th
• Midterm, 25%, Date: March 1st 8:00-9:30am in-class
• Final, 35%, Date: TBA

Learning Activities

Assignments will be assigned approximately biweekly. Each assignment should be prepared using handwritten answers, and RMarkdown or LaTeX using knitR. The solutions (including both written answers converted to PDF files and R markdown files) need to be submitted to Canvas before the deadlines. It would be best if you started on assignments as soon as material is covered. Do not wait until the last minute. Assignments can be submitted in groups of up to 3. Assignments should be printed and submitted in-class. Tentative due dates are January 23rd, February 6th, February 27th, March 13th, March 27th, April 13th.

The group project will test your ability to apply the knowledge of this course to real-world data as well as computing skills. The minimum group size is 2, while the maximum group size is 3. Use R to analyze some given data set, or choose your own. A one- to two-page proposal (worth 5% of your final mark) describing your chosen dataset, scientific questions, and regression techniques used to answer those questions is required and should be submitted to Canvas by 11:59 pm sharp on Monday, March 6th 2023. A full report (worth 25% of your final mark) describing the project methodology, results and conclusions, together with all the R code used, is required and should be submitted to Canvas by 11:59 pm sharp on Monday, April 13th 2023. Further detailed instructions will be given in-class on Wednesday, February 1st 2023.

Course schedule

Weeks	Topics covered
1	Course overview, Introduction to regression concepts, Simple linear regression

2	Matrix algebra review, Normal theory for regression, Matrix algebra applied to simple linear regression
3	Model adequacy, violating linear model assumptions, Residual analysis, Leverage and influence
4	Model diagnostics through residuals, variable transformations
5	Variable selection methods, Multicollinearity
6	Variance inflation factors, Ridge and LASSO regression
7	Regression with indicator and categorical variables Midterm exam on Wednesday, March 1st In-class
8	Theory of likelihood, Exponential family of distributions, Likelihood ratio tests
9	Generalized linear models, generalized and iteratively re-weighted least squares,
10	Logistic regression, Residual deviance
11	Poisson regression, Overdispersion, Goodness of fit test, Negative binomial regression
12	Spline regression, Generalized additive models
13	Mixed effects models, Review

Late policy

Except for extreme situations (e.g., illness, childbirth, or bereavement), the following policy is applied to late assignments or project:

• 0 to 24 hours late: 25%-mark deduction (e.g., if an assignment is worth 20 marks, then 5 marks will be deducted from the assignment mark; no negative marks will be given.).
• 24 to 48 hours late: 50%-mark deduction
• More than 48 hours: no mark.

Missed exam policy

Students who, because of unforeseen events, are absent during the term and are unable to complete tests or other graded work should generally discuss with their instructors how they can make up for missed work, according to written guidelines given to them at the start of the course (see Grading Practices). Instructors are not required to make allowance for missed tests or incomplete work not satisfactorily accounted for. If ill-health is an issue, students are encouraged to seek attention from a health professional. Campus Health and Counselling will usually provide the documentation only to students who have been seen previously at these offices for treatment or counselling specific to conditions associated with their academic difficulties. Students who feel that requests for consideration have not been dealt with fairly by their instructors may take their concerns first to the Head of the discipline and, if not resolved, to the Office of the Dean. Further information can be found at:
http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,48,0,0.

There will be no make-up midterm exams. If the absence is satisfactory, the student's final exam will be worth more in the final grade.

Passing/Grading Criteria

To pass this course, students must achieve a passing grade (greater than or equal to 50%) on the final examination. If they do not, their final examination mark will account for their entire grade.

Learning Materials

The course notes will be the main resource of material for this course, and supplementary readings will be provided from the following resources:
• Douglas C. Montgomery, Elizabeth A. Peck (Author), and G. Geoffrey Vining, Introduction to Linear Regression Analysis, (Fifth Edition), 2012. Wiley & Sons, Inc. https://go.exlibris.link/1krrFXSn
• Simon Wood, Generalized Additive Models: An Introduction with R. 2nd Edition, 2017. CRC-Chapman & Hall.

https://go.exlibris.link/CZTkSPxy
• W. John Braun, Regression and Generalized Linear Models, 2021. On Canvas.

For learning R and RMarkdown, some great resources are:
• W. John Braun and Duncan J. Murdoch, A first course in statistical programming with R, (Second Edition), 2016. Cambridge University Press. [An excellent introduction to the basics of R.]
• Hadley Wickham, Advanced R, (Second Edition), 2019. CRC press. [More detail to R than you could ask for, from a programming perspective.]
• Garrett Grolemund and Hadley Wickham, R for Data Science, 2018. O'Rielly Media Inc. [An excellent guide to modern data analysis with R that uses the tidyverse package.]
• Yihui Xie, J.J. Allaire, and Garrett Grolemund, R Markdown: The Definitive Guide, 2020. CRC Press. [All you want to know about RMarkdown used for assignments and projects.]

Other Course Policies

Academic Integrity

The academic enterprise is founded on honesty, civility, and integrity. As members of this enterprise, all students are expected to know, understand, and follow the codes of conduct regarding academic integrity. At the most basic level, this means submitting only original work done by you and acknowledging all sources of information or ideas and attributing them to others as required. This also means you should not cheat, copy, or mislead others about what is your work. Violations of academic integrity (i.e., misconduct) lead to the breakdown of the academic enterprise, and therefore serious consequences arise and harsh sanctions are imposed. **For example, incidences of plagiarism or cheating usually result in a failing grade or mark of zero on the assignment or in the course.** Careful records are kept to monitor and prevent recidivism.

A more detailed description of academic integrity, including the University's policies and procedures, may be found in the Academic Calendar at: http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,54,111,0

Final Examinations

You can find the Senate-approved term and examination dates here. Except in the case of examination clashes and hardships (three or more formal examinations scheduled within a 27-hour period) or unforeseen events, students will be permitted to apply for out-of-time final examinations only if they are representing the University, the province, or the country in a competition or performance; serving in the Canadian military; observing a religious rite; working to support themselves or their family; or caring for a family member. Unforeseen events include (but may not be limited to) the following: ill health or other personal challenges that arise during a term and changes in the requirements of an ongoing job.

Further information on Academic Concession can be found under Policies and Regulation in the Okanagan Academic Calendar http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,48,0,0

Grading Practices

Faculties, departments, and schools reserve the right to scale grades in order to maintain equity among sections and conformity to University, faculty, department, or school norms. Students should therefore note that an unofficial grade given by an instructor might be changed by the faculty, department, or school. Grades are not official until they appear on a student's academic record.

http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,41,90,1014

Resources to Support Student Success:

UBC Okanagan Disability Resource Centre

The DRC facilitates disability-related accommodations and programming initiatives to remove barriers for students with disabilities and ongoing medical conditions. If you require academic accommodations to achieve the objectives of a course please contact the DRC at:

UNC 215 250.807.8053
Email: drc.questions@ubc.ca
Web: www.students.ok.ubc.ca/drc

UBC Okanagan Equity and Inclusion Office

Through leadership, vision, and collaborative action, the Equity & Inclusion Office (EIO) develops action strategies in support of efforts to embed equity and inclusion in the daily operations across the campus. The EIO provides education and training from cultivating respectful, inclusive spaces and communities to understanding unconscious/implicit bias and its operation within in campus environments. UBC Policy 3 prohibits discrimination and harassment on the basis of BC's Human Rights Code. If you require assistance related to an issue of equity, educational programs, discrimination or harassment please contact the EIO.

UNC 325H 250.807.9291
Email: equity.ubco@ubc.ca
Web: www.equity.ok.ubc.ca

Student Wellness

At UBC Okanagan health services to students are provided by Student Wellness. Nurses, physicians and counsellors provide health care and counselling related to physical health, emotional/mental health and sexual/reproductive health concerns. As well, health promotion, education and research activities are provided to the campus community. If you require assistance with your health, please contact Student Wellness for more information or to book an appointment.

UNC 337 250.807.9270
Email: healthwellness.okanagan@ubc.ca
Web: www.students.ok.ubc.ca/health-wellness

Office of the Ombudperson

The Office of the Ombudsperson for Students is an independent, confidential and impartial resource to ensure students are treated fairly. The Ombuds Office helps students navigate campus-related fairness concerns. They work with UBC community members individually and at the systemic level to ensure students are treated fairly and can learn, work and live in a fair, equitable and respectful environment. Ombuds helps students gain clarity on UBC policies and procedures, explore options, identify next steps, recommend resources, plan strategies and receive objective feedback to promote constructive problem solving. If you require assistance, please feel free to reach out for more information or to arrange an appointment.

UNC 328 250.807.9818
Email: ombuds.office.ok@ubc.ca
Web: www.ombudsoffice.ubc.ca

Student Learning Hub

The Student Learning Hub is your go-to resource for free math, science, writing, and language learning support. The Hub welcomes undergraduate students from all disciplines and year levels to access a range of supports that include **tutoring in math, sciences, languages, and writing, as well as help with study skills and learning strategies**. Students are encouraged to visit often and early to build the skills, strategies and behaviors that are essential to being a confident and independent learner. For more information, please visit the Hub's website.

LIB 237 250.807.8491
Email: learning.hub@ubc.ca
Web: www.students.ok.ubc.ca/slh

The Global Engagement Office

The Global Engagement Office provides advising and resources to assist International students in navigating immigration, health insurance, and settlement matters, as well as opportunities for intercultural learning, and resources for Go Global experiences available to all UBC Okanagan students, and more.

Come and see us – we are here to help! You may also contact geo.ubco@ubc.ca

Safewalk
*Don't want to walk alone at night? Not too sure how to get somewhere on campus? Call Safewalk at **250-807-8076.***
For more information, see: www.security.ok.ubc.ca