Statistics 6450: Applied Regression Analysis
Spring 2018 Course Syllabus

Instructor: Dr. Subhadeep Paul

Lectures: 3:00 pm – 4:50 pm on Wednesdays and Fridays in Denney Hall - Room 253.

Office Hours: 1:30 pm - 2:30 pm on Wednesdays and Thursdays at 231 Cockins Hall. You are welcome to walk in anytime during the office hours. However, if you can't make it during the specified hours, you can email me to request an appointment at a different time. Please note such appointments will not be available on short notice.

Email: paul.963@osu.edu.

It is usually difficult to effectively communicate mathematical/statistical concepts through email. So please consider whether your question would be best answered in person in class or during office hours. You can, however ask a question through email for a discussion in the next class (See paragraph on Questions below). I will try my best to be prompt in responding to emails, however a quick response is not guaranteed.
To protect your privacy, all course email correspondence must be conducted using your valid OSU name.# email account.

Office: 231 Cockins Hall

Grader: Yiyang Lin (lin.1459@osu.edu)

Grader Tutor room hours: TBD

Course Website: Important announcements, course materials and lecture notes, homework problems, computing references, and other information about the class will be posted on Carmen (carmen.osu.edu, login with your web ID).

Textbook: The required textbook for this course is:

- *Applied Linear Regression Models*, 4th edition, by Kutner, Nachtsheim, and Neter.

Also see the section on additional references and R resources.

Course Description: Statistics 6450 is intended to be an introduction to regression analysis techniques. Its focus will be on the application of linear regression models in practice but will also cover basic theory of the linear model. Topics of Stat 6450 include:

- Simple Linear Regression (SLR) model
  - Methodology for fitting models
  - Statistical inference
  - Diagnostics for verification of assumptions and their remedies
  - Solving using matrix algebra
- Multiple Linear Regression (MLR) model
  - Methodology for fitting models
  - Statistical inference
  - Binary indicator (1/0) and qualitative predictors
  - Diagnostic measures of model fit
  - Variable selection and model building
- Other Models
  - Generalized Linear Models (GLMs)
  - Logistic regression
  - Regression with ordinal and nominal polytomous response

Prerequisites / Co-requisites:
Statistics 6201 or equivalent.

Exclusions:
Not open to students with credit for Stat 645 (Stat 6450 under semesters)

Course Objectives:
By the end of the course, you should:

- Understand the motivation of regression analysis
- Understand the theoretical assumptions behind the linear model and their importance in properly conducting a regression analysis
- Know how to estimate the parameters in regression models
- Be able to validate the modeling assumptions with formal tests and visual diagnostic tools
- Know how to make inferences regarding the linear model
- Be able to build and validate regression models in a principled manner
- Be able to apply the above knowledge and techniques in on your own data or problems

Course requirements: You are responsible for all material covered in class and in the required readings; this includes derivation, proofs, computational techniques, etc. This is an applied course and the emphasis will be on applying concepts learned in class to real-world datasets. However, there will be an emphasis on theoretical concepts which will help you better understand and apply the techniques covered in class. You are expected to be comfortable with mathematical statistics, basic matrix operations and linear algebra. There are computational elements to this class; you will learn how to use software to analyze data and apply concepts learned during the lectures.

Course Materials: A reading list from the required textbook will be provided for each lecture. I will also post the lecture notes in Carmen after each class.

Questions: Asking questions is one of the best ways to learn a new material. I will set aside a few minutes in each class for a discussion on any question you have on the entire course material covered up to that point. This is of-course in addition to any clarifying question during the lecture. If you would prefer to ask a question anonymously, you can send me an email with your question and I will discuss that in the next class.

Homework Assignments: Homework will be assigned regularly. It will consist of both written problems and small computer programming/data analysis problems. You are encouraged to work together on the problems, but each student must hand in their own work, written in their own words. Do not copy any part of another student's homework including computer output. Use of homework solutions distributed in previous offerings of the course or available on the web constitutes academic misconduct and will be handled according to university rules. A hard copy of the homework solutions should be submitted at the beginning of class on the due date. The written solutions may be handwritten or typed. Please be sure that the questions are clearly labeled, all supporting work (including computer code) can be easily identified, and that all figures/tables are referenced and interpreted in the text. Electronic versions of homework solutions will not be accepted unless permission from the instructor is obtained in advance.

Exams: There will be an in-class midterm exam and a final exam. Coverage includes lecture material, assigned reading, and homework. Tentative dates are provided on the weekly lesson plan. Statistical tables will be provided as needed. Calculators may be used, but no communication devices are allowed (e.g. mobile phones). Makeup exams require a valid excuse and official proof if I am notified in advance or as soon as possible. A make-up exam must be taken within a week of the missed exam. Exceptions to this policy are permitted only in extreme situations such as serious injury immediately prior to an exam or severe illness requiring hospitalization.

Project: A data-analysis class project will consist of a combination of a written portion and oral presentation, and will require use of the R software. Suggested project topics will be provided in class.

Computing: We will be using the `R` statistical computing software. `R` may be downloaded for free from http://www.r-project.org/. Many students prefer to use the interface RStudio, available for free at http://www.rstudio.com. Students are welcome to use other software for assignments (such as SAS, SPSS, minitab, etc), however `R` is strongly encouraged.

Attendance: Regular attendance and class participation is required. To comply with Federal title IV regulations, an actual roll call to verify attendance will be conducted during the first 2 lectures.

Grading: In order to obtain full credit on homework and exam problems you need to show your justification for or full work. Answers without work will not receive full credit. The following is a breakdown of the final course grade:

Homework: 30%
Midterm Exam: 20%
Final Exam: 30%
Project: 20%

The following rubric will be used to compute the final letter grade: A: 93 – 100, A-: 90 – 92.9, B+: 87– 89.9: B: 83 – 86.9, B-: 80-82.9, C+: 77-79.9, C: 73 – 76.9, C-: 70–72.9, D+: 67 – 69.9, D: 60–66.9, E: below 60.

***Optional* References:**

- Related Texts
  - *Applied Regression Analysis*, Wiley. Normal Draper and Harry Smith (1998).
  - *Introduction to Linear Regression Analysis*, Wiley-Interscience. Douglas Montgomery (2006).
- Advanced Regression
  - *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2nd ed. 2015 edition. Frank E. Harrell, Jr. (Chapters 1, 2, 4, 6, 9, 10, 11 contain a more advanced coverage of techniques in regression modeling with particular emphasis towards applications and a focus on problem solving *strategies*. I think this is a wonderful book and while it may not be needed in this class, I have found it to be useful on many occasions.)
  - *Linear Regression Analysis*, 2nd Edition. George A. F. Seber, Alan J. Lee. (For a more advanced coverage of theoretical issues of the linear model [for the curious]. This text also includes some relevant background on linear algebra in the appendix)
- R Resources
  - *Using R for Data Analysis and Graphics*. J.H. Maindonald. This text if available for free online: `https://cran.r-project.org/doc/contrib/usingR.pdf`
  - *R for Beginners*. Emmanuel Paradis. This text is available for free online: `https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf`
  - *R Bootcamp*. Jared Knowles. This is an online-based R bootcamp: `https://www.jaredknowles.com/r-bootcamp/`
  - *Quick R*. This is a good online reference for R topics `http://www.statmethods.net/`
  - *Advanced R*. Hadley Wickham. This text is available for free online: `http://adv-r.had.co.nz/` (*Advanced R* is for *advanced* users of `R` but is a very good `R` resource)

Special Considerations: If a situation exists or arises that you think may hinder your progress in this class, you must notify me as soon as possible.

Academic Misconduct: It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Disability Services: Students with disabilities that have been certified by the Office for Disability Services will be appropriately accommodated and should inform the instructor as soon as possible of their needs. The Office for Disability Services is located in 150 Pomerene Hall, 1760 Neil Avenue; telephone 292-3307, TDD 292-0901; http://www.ods.ohio-state.edu/.