16:960:563:01 Index 16671 – Regression Analysis (Fall/2023)

Meeting Time: Thursday (6:00-9:00 pm) in ARC-107

Instructor: Jack Mardekian, PhD. Retired (1/ 2020) from Pfizer Inc

Teaching Assistant: No TA this semester.

Contact information: Rutgers email: mardekia@stat.rutgers.edu (email is the BEST OPTION to contact me – I respond quickly to e-mails. Please do not use the Canvas Inbox Folder, I do not monitor the Inbox Folder on a regular basis.)

908.304.8400 (mobile, to be used only for a course-related emergency), 848.445.7647 (office)

Student Support Hours:
1. Thursday, 3:30-4:30 pm in my office, Hill Center Rm 465
2. Friday, 9-10 am, Zoom: https://rutgers.zoom.us/j/7358396947?pwd=N1BHMFj72ErGQ5RkFvWDVjakpaZz09
3. Before and after class (I will be outside classroom by 5:30 pm)
4. Support hours by appointment are available.

TA Support Hours: Not applicable.

Course Description (from Course Catalog):
Prerequisite: Level IV Statistics.
Review of basic statistical theory and matrix algebra; general regression models, computer application to regression techniques, residual analysis, selection of regression models, response surface methodology, nonlinear regression models, experimental design models, analysis of covariance. Emphasis on applications.

Levels IV and V Statistics: The Level V Statistics prerequisite for some courses may be fulfilled by 16:960:563 or 586 or 593, while the Level IV Statistics prerequisite may be fulfilled by 01:960:401 or 01:960:484 or 16:960:590 or Level V Statistics.

Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining

First Day Course Materials (a module in Canvas): Wiley arranged a discounted price of $106.25 (from $124.00 at the online textbook site) for an eBook version of the textbook. Please be sure to opt out from the offer in Canvas by September 19 if you arrange to access the textbook elsewhere.

Material in the textbook will be supplemented. I placed a copy of the 6th edition on Reserve in the Library of Science and Medicine. FYI - the Rutgers Library does not have access to the electronic version of the 6th edition. The Rutgers Library has a copy of the 5th Edition that can be reviewed online.

Short, asynchronous lectures may be recorded as needed and placed into Canvas.

Canvas class site: 2023FA - REGRESSION ANALYSIS 16:960:563:01 (instructure.com)
Class notes, assignments, and announcements will be posted on this site.

Technical/Technology Requirements:
Software Considerations
- R Statistical Software, which is powerful and free, is required for this course.
  - The link to the current version of the software is provided here: https://cloud.r-project.org/, If you have a previous version that is at least R version 4.0.0 (2020-04-24) then you do not need to install the current version.
- RStudio is a free platform for both writing and running R, available at www.rstudio.org. Some students find it friendlier than basic R.
- The learning curve for R/RStudio is very steep. Students can become proficient in a few weeks. Manuals are very helpful to learn R. Manuals are provided here: http://cran.r-project.org/manuals.html.
• Limited software instruction, in-class demonstration, and code to accompany lectures and assignments will be provided during class meetings. Basic computer skills are assumed. Typical to any software language, R is only learned by writing, executing, and fixing your own code. Data management/wrangling will be kept to a minimum.

• You should install R as soon as possible and familiarize yourself with basic operations. You will need to be able to refer to R functions/library manuals and examples that are available through the Internet to successfully complete homework assignments. Students must take the initiative in seeking advice and help from appropriate sources including documentation resources, other students, the Teaching Assistant, Instructor. We will also be reviewing SAS output, but you will not be required to execute SAS (SAS is no longer listed as a standard offering by Rutgers Computer Labs).

• Rutgers Office of Information Technology website: https://it.rutgers.edu/new-brunswick/

• Rutgers Student Tech Guide: see https://it.rutgers.edu/technology-guide/students/#new-brunswick.

• If you do not have the appropriate technology for financial reasons, please email the Dean of Students at deanofstudents@echo.rutgers.edu for assistance. If you are facing other financial hardships, please visit the Office of Financial Aid: https://financialaid.rutgers.edu/.

• Teaching and Learning with Technology web page: https://tlt.rutgers.edu/

• Technology resources for students: https://coronavirus.rutgers.edu/technology-resources-for-students/

Course Learning Goals: Students will gain an understanding of regression methods including simple linear regression, least squares estimation, multiple linear regression, checking model adequacy, diagnostics for leverage and influence, variable selection and model building, multicollinearity, weighted least squares, and logistic regression. Students will be able to execute regression procedures using R software and perform analyses and interpret the resulting computer output. Students will be able to perform analyses based on SAS and Minitab output presented in the required textbook for the course. Programming in SAS/Minitab is out of scope for the course.

Grading: Attendance and successful completion of HW assignments are two basic requirements for passing this course.

90+ = A; 87-89 = B+; 80-86 = B; 77-79 = C+; 70-76 = C; 60-69 = D; < 60 = F.

If the highest average among students is < 90, then a grading scale that is curved may be applied. Grades will be assigned by relative ranking within each section. Students with poor attendance (defined by excessive unexcused absences and/or failure to watch asynchronous lectures) may not take part in the curved grading scale.

In-class Midterm Exam (35%) 26Oct2023
In-class Final Exam on material since Midterm Exam  (35%) 14Dec2023
One data analysis project (15%)
HW assignments and class participation (15%)

Schedule of Topics: There are 14 class meetings + Final Exam. The first class meeting is 9/7 and the last class meeting is 12/7.

We will closely follow the textbook material using slides (available in Canvas) generated by the Authors of the text.

Weeks 1-2
Appendix E - Introduction to R to Perform Linear Regression Analysis
Chapter 1.3 Uses of Regression, Chapter 2 Simple Linear Regression Model, Least-Squares Estimation of the Parameters, Hypothesis Testing, Confidence Intervals, Prediction Intervals, Estimation by Maximum Likelihood

Weeks 3-4
Chapter 3. Multiple Regression Models
Matrix approach to simple linear model, hypothesis tests, Type 1 and 3 SS, extra SS, simultaneous inferences, partial F test (reduced versus full model), multicollinearity

Weeks 5-6
Chapter 4. Model Adequacy Checking
Residual analysis, residual plots, PRESS statistic
Chapter 6. Diagnostics for Leverage and Influence
Leverage, Influence, Measures of Influence

Week 7
Chapter 7. Polynomial Regression Models
7.2 Polynomial Models in One Variable, 7.4 Polynomial Models in Two or More Variables
Cross Validation
Chapter 8. Indicator Variables
8.1 General Concept of Indicator Variables
Chapter 9. Multicollinearity
9.3 Effects of Multicollinearity, 9.4 Multicollinearity Diagnostics, 9.5.3 Ridge Regression

Week 8: Midterm Exam
Weeks 9-10
Chapter 10. Variable Selection and Model Building
10.1.3 Criteria for Evaluating Subset Regression Models, 10.2.1 All Possible Regressions, 10.2.2 Stepwise Regression Methods
Lasso Regression

Weeks 11-12
Chapter 13. Generalized Linear Models
13.2 Logistic Regression Models, Estimating the Parameters in a Logistic Regression Model, 13.2.3 Interpretation of the Parameters in a Logistic Regression Model, Statistical Inference on Model Parameters, 13.2.5 Diagnostic Checking in Logistic Regression

Weeks 13-14
15.4.1 Bootstrap Sampling in Regression, 15.4.2 Bootstrap Confidence Intervals
Chapter 14. Regression Analysis of Time Series Data (may be omitted if there is not enough time)

Final Exam

Other information: Appropriate use of ChatGPT and other AI applications is allowed with proper attribution. Students should cite any AI-generated material that informed their work (e.g., homework, projects).

Policies (see https://sasundergrad.rutgers.edu/degree-requirements/policies) Many undergraduate policies are applicable to graduate students.

- **Attendance:** Attendance is expected and monitored. If you will miss or have missed a class, please use the University absence reporting system to indicate the date for your absence: https://sims.rutgers.edu/ssra/. An email is automatically sent to me. There is no need to provide supporting documentation for your absence. Follow the guidance of the University absence reporting system for longer periods of absence. I am available to collaborate with you on the missed material.

  Please do not attend class if you are ill, especially if you have symptoms of Covid (see https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html).

- **Disability Accommodations:** To receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation. For further information consult the policies and procedures of the Office of Disability Services website: https://ods.rutgers.edu.

- **Academic Integrity:** Rutgers University takes academic dishonesty very seriously. By enrolling in this course, you assume responsibility for familiarizing yourself with the Academic Integrity Policy and the possible penalties (including suspension and expulsion) for violating the policy. http://nbacademicintegrity.rutgers.edu/ As per the policy, all suspected violations will be reported to the Office of Student Conduct. Please consult with me if you are ever in doubt regarding Academic Integrity.
Student Support and Mental Wellness: Here is a list of Rutgers resources supporting students in their academic success and mental wellness,

- Student Success Essentials: [https://success.rutgers.edu/](https://success.rutgers.edu/)
- Student Support Services: [https://www.rutgers.edu/academics/student-support](https://www.rutgers.edu/academics/student-support)
- The Learning Centers: [https://rlc.rutgers.edu/](https://rlc.rutgers.edu/)
- The Writing Centers (including Tutoring and Writing Coaching): [https://writingctr.rutgers.edu/](https://writingctr.rutgers.edu/)
- Rutgers Libraries: [https://www.libraries.rutgers.edu/](https://www.libraries.rutgers.edu/)
- Office of Veteran and Military Programs and Services: [https://veterans.rutgers.edu/](https://veterans.rutgers.edu/)
- Student Health Services: [http://health.rutgers.edu/](http://health.rutgers.edu/)
- Office for Violence Prevention and Victim Assistance: [www.vpva.rutgers.edu/](http://www.vpva.rutgers.edu/)

Useful Statistics web sites (accessed on August 25, 2023)
- [https://online.stat.psu.edu/stat501/](https://online.stat.psu.edu/stat501/) Course notes for STAT 501: Regression Methods taught by Penn State’s Department of Statistics. These notes are offered as open educational resources. These notes are free to be used under Creative Commons license [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/)
- [http://stattrek.com/](http://stattrek.com/)
- [http://r4stats.com/](http://r4stats.com/)
- [Big Book of R](https://www.ferranti.net/rbook/) compiled by Oscar Baruffa (See Section 30: Statistics)

Learn R in 39 minutes Andrew Gard
[https://www.youtube.com/watch?v=yZ0bV2Afkjc](https://www.youtube.com/watch?v=yZ0bV2Afkjc)

Directory of Andrew Gard videos
[https://www.youtube.com/c/EquitableEquations/videos](https://www.youtube.com/c/EquitableEquations/videos)

Markdown by Andrew Gard
[https://www.youtube.com/watch?v=asHhuHRxhvo](https://www.youtube.com/watch?v=asHhuHRxhvo)