STA671: REGRESSION AND CORRELATION
Spring 2011

Lectures:

Time: 11:00-11:50am, MWF (January 12 – March 2)
Place: Whitehall Classroom Building Room 238

Computing labs:

Section 001: M 5:00-6:15pm (January 24 – February 28) in WCB Rm 307
Section 002: W 5:00-6:15pm (January 19 – February 23) in WCB Rm 307
Section 003: T 8:00-9:15am (January 18 – February 22) in WCB Rm 309

Course web-site:

Course materials will be posted via blackboard. Details will be announced soon.

Instructor

Dr. Simon Bonner
Office: 839 Patterson Office Tower
Phone: 859-257-4950
E-mail: simon.bonner@uky.edu
Office hours: Monday and Thursday 4:00-5:00 or by appointment

Teaching Assistants

Sections 001 & 002
Zilong Wang
Office: 859 Patterson Office Tower
E-mail: zilong.wang@uky.edu
Office hours: TBA

Section 003
Qian Fan
Office: 818 Patterson Office Tower
E-mail: qfa222@uky.edu
Office hours: TBA

Description

How does annual rainfall affect crop yields? Is there a link between smoking and the incidence of cancer? What are the best predictors of a child's success in school?

Regression and correlation serve as the basis for studying relationships between different variables and for answering questions about associations which arise in all areas of quantitative research. This course will introduce the essential methods of both simple and multiple linear regression. Our attention will focus on developing a non-mathematical understanding of the concepts behind these methods, learning to choose between methods, implementing these techniques in standard statistical software, and interpreting and summarizing the results. Specific content will include:

- computing correlations to assess the degree of association between two variables,
- fitting simple and multiple regression models,

- interpreting the results of these procedures,
- checking that modelling assumptions are satisfied,
- correcting problems that might arise, and
- selecting the best model/set of variables to predict an outcome of interest.

Some extensions of the basic methods will be introduced if time permits.

Course Objectives

The primary goals of this course are that you develop the skills necessary:

1) to perform a full regression analysis of simple data using the SAS statistical software,
2) to interpret and summarize the results you obtain, and
3) to think critically about the use of linear regression methods.

Course activities will help you to develop a conceptual understanding of correlation and regression analysis, to implement these methods in the SAS statistical software package and to interpret the results, and to evaluate the use of these methods in your own work and in published literature in your field.

Evaluation

Your performance in the course will be assessed through the following:

1. Assignments (50% of final grade):
   There will be a total of 6 weekly assignments based around the weekly computer labs that will require you to implement the methods you have learned in the course and to answer questions and/or submit short reports on your work. Tasks will range from simple implementation of the methods and interpretation of results to broader problems requiring you to demonstrate a general principal or to select between methods to answer a specific question.
   Assignments are to be submitted at the **start** of the following computer lab.

2. Final data analysis (30% of final grade):
   The final data analysis will require you to apply the methods you have learned to a real data set. You will be required to submit a report describing your analysis, justifying the methods you have chosen, and interpreting the results. More details will be provided closer to the end of the course.

3. Final exam (20% of final grade):
   There will be one exam on the final day of classes, March 2. This exam will cover all material introduced during the lectures, labs, and on assignments. Make-up exams will only be provided for students with an approved, official university excuse.

Note: All work to be assessed must be submitted either in hard copy or by email, as specified, by the assigned due dates. Late work will only be accepted with an official university excuse. Please the the UK policy on Student Rights and Responsibilities, Part II, Section 5.2.4.2 (http://www.uky.edu/StudentAffairs/Code/part2.html) for details on official excuses.

Materials

The recommended textbook for this course is:

Dielman, T.E. (2004) *Applied Regression Analysis: A Second Course in Business and Economic Statistics*, 4th Edition. Belmont, CA: Brooks/Cole Thomson Learning.

You may also find useful information in other introductory regression analysis texts. Come to my office hours if you would like recommendations on further reading.

Policies

Academic accommodations:
Please see me as soon as possible if you have a documented disability that requires academic accommodations. In order to receive accommodations in this course, you must provide me with a Letter of Accommodation from the Disability Resource Center (Room 2, Alumni Gym, 257-2754, jkarnes@uky.edu) for coordination of campus disability services available to students. We can then work together to find the best solutions for you.

Attendance:
Your attendance in class will not be checked but regular attendance is the best way to learn the material. You are responsible for all material covered during the class. If you know that you will be away ahead of time then I recommend stopping by my office hours and I will do the best I can to accommodate.

Academic integrity:
Academic dishonesty of any form will not be tolerated. Minor forms of dishonesty will result in a score of zero for that component of the course work. More serious forms of dishonesty will be reported to the university. Further information on plagiarism, how to avoid plagiarism, and the university's academic offense policy are available from the web-site of the Office of Academic Ombud Services (http://www.uky.edu/Ombud).

Classroom behaviour:
I view my role as a teacher is not to deliver knowledge to you, but to guide you through your own learning process. The course will require participation from every student, and I expect you to come to class willing to participate by completing exercises to the best of your ability and by engaging in class discussions. I will work to maintain an open and respectful atmosphere in the classroom and expect you to do the same. This includes providing others with the opportunity to ask questions and express opinions. Please turn of all mobile devices before class so that they will not interrupt the other students.

Timetable (tentative)

Topics	Expected Timing
Association between Two Variables • Scatterplots • Correlation	Jan 12 – 14
Simple Linear Regression (Chapter 3) • Least squares estimation • Confidence intervals • Significance tests • ANOVA tables • Prediction	Jan 19 – 28
Multiple Linear Regression (Chapters 4, 5, and 7) • Estimation with several predictors • Confidence intervals • Significance tests • ANOVA tables • Prediction • Multi-collinearity • Interactions • Indicator variables • Polynomial regression and transformation of variables	Jan 31 – Feb 11
Assessing Regression Assumptions (Chapter 6) • Assumptions of a regression model • Residuals • Checking assumptions and fixing problems • Detecting outliers	Feb 14 – Feb 21
Model Comparison and Model Selection (Chapter 8) • All possible regressions • Tests for comparing two models • Stepwise model selection • Criterion based model selection	Feb 23 – Feb 28
Review and extra topics as time allows	Mar 2 – Mar 4

Chapter information refers to the recommended text, Dielman (2004). Note that the course will not cover all of the material in these chapters and will also contain material not included in this text.