As a boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together - we are Purdue. — The Purdue Honors Pledge

Instructor: Dr. Dabao Zhang
Office: 534 Mathematical Sciences Building
E-mail: zhangdb@stat.purdue.edu

Office Hours: T 1:20-2:50PM

Course Web Page: http://www.stat.purdue.edu/~zhangdb/stat526/

Prerequisite: STAT 525 and 517 (or STAT 525 and 528). You must have the working knowledge of basic statistical inference and modeling (e.g., maximum likelihood estimate, likelihood ratio test, and linear regression models).

Course Objectives: To understand the concepts and methods in generalized linear models, survival analysis, (generalized) linear mixed models and GEE methods (nonparametric regression and classification and regression trees may be covered if we have enough time). To apply these methods properly to real data using R, draw valid conclusions, and present the results concisely.

Computer Software: We will be using R, an open-source environment which implements the S language (also the basis of the S-PLUS systems), as the primary platform for data manipulation, computation and graphical display. Computer software is NOT a substitute for understanding the statistical methods, and you will not have access to a computer during exams. R can be downloaded from CRAN (https://cran.r-project.org/), the Comprehensive R Archive Network. The tutorial document, An Introduction to R (http://cran.r-project.org/doc/manuals/R-intro.pdf), should be helpful to you, especially if you had little previous exposure to R/S/Splus.

Mailing List: A mailing list has been arranged for this course. I will send e-mail to this list with any special announcements or reminders.

Textbook: Extending the Linear Model with R by J. J. Faraway

References:


Course Outline:

- Review on Linear Models/R
- Generalized Linear Models
- Categorical Data Analysis
- Survival Data Analysis
- Linear Mixed Models
- Clustered/Repeated Categorical Data
- Nonparametric Regression
- Classification and Regression Trees (Optional)
- Neural Network (Optional)
- Project Presentation

**Homework:** There will be about 10-12 homework assignments during the semester. Everybody must hand in his (her) own solutions but a group discussion is encouraged. *Any overdue homework will not receive credit.* When turning in your homework, each problem must include all relevant graphs and tables. *Do not include irrelevant graphs and tables!* Your solutions must be easily readable and appropriately labeled. Any graph or table that is presented without comments will be ignored.

**Project:** At the end of the semester there will be a project where groups of about 3-5 students find a real-world problem to analyze and summarize. A month before the end of the semester, each group will submit a project proposal. The proposal will describe group members, the scientific question to be addressed, the available data, and statistical analysis methods that will be used. I encourage you to talk to me before submitting the proposal. Each group will write a final report that will be due a week before the end of the class. During the last week of class each class member will receive a report from another group, and evaluate the analysis and conclusions of the project.

**Re-grades:** Since the professor and grader are fallible human beings, occasionally errors will occur in grading. For this reason, students are able to request that such an error be corrected. Two types of error can occur. A type I error occurs if points are deducted for a correct solution. A type II error occurs if sufficient points are not deducted for an incorrect solution. Any request for a re-grade must be made in writing with a detailed explanation of the suspected error ("Please look at problem 4" is not considered a detailed explanation) and submitted within a week, or it will be ignored. Please note that a re-grade request is different from the questions “Can you help me figure out what I did wrong here?”, or “I don’t understand the posted solutions”, which are entirely appropriate for office hours.

**Final Grade:** The final grade is broken down into Homework 20%, Evening Exam I 30%, Evening Exam II 30%, and Project 20%. Your final grade will be determined by taking your received percentage to the following scale: 99+ = A+, [90,99) = A, [80,90) = B, [70,80) = C, [55,70) = D, [0,55) = F. The minimum score needed for a given letter grade could be lowered if necessary but will not be raised. The project will be ongoing throughout the semester. It will be a group project involving a written summary and a class presentation in the last week of class.

**Attendance:** *Attendance is mandatory.* I will pick around 10 random days to do roll call per semester (some times this is replaced by homework distributing). If you miss more than 3 lectures without official excuse, you will start to loose points in your cumulative grade. Each absence in addition to 3 will lead to losing 1 point.

Whenever you miss a class, you are responsible for announcements and the material covered during lectures. If you cannot attend an exam at the assigned time, notice must be given at least a week prior to
the exam in order to decide on a different (preferably earlier) time.

**Academic Integrity Statement:** Any test, paper or report submitted by you and that bears your name is presumed to be your own original work that has not previously been submitted for credit in any other course unless you obtain prior written approval to do so from your instructor. In all of your assignments, including your homework or drafts of papers, you may use words or ideas written by other individuals in publications, web sites, or other sources, but only with proper attribution. “Proper attribution” means that you have fully identified the original source and extent of your use of the words or ideas of others that you reproduce in your work for this course, usually in the form of a footnote or parenthesis.

As a general rule, if you are citing from a published source or from a web site and the quotation is short (up to a sentence or two) place it in quotation marks; if you employ a longer passage from a publication or web site, please indent it and use single spacing. In both cases, be sure to cite the original source in a footnote or in parentheses. If you are not clear about the expectations for completing an assignment or taking a test or examination, be sure to seek clarification from your instructor beforehand.

Finally, you should keep in mind that as a member of the campus community, you are expected to demonstrate integrity in all of your academic endeavors and will be evaluated on your own merits. So be proud of your academic accomplishments and help to protect and promote academic integrity at Purdue. The consequences of cheating and academic dishonesty - including a formal discipline file, possible loss of future internship, scholarship, or employment opportunities, and denial of admission to graduate school - are simply not worth it.

**In the Event of a Major Campus Emergency:** Course requirements, deadlines and grading percentages are subject to change that may be necessitated by a revised semester calendar or other circumstances beyond the instructor’s control. Here are the ways to get information about changes in this course: course web page, instructor’s email, and instructor’s phone.

**Important Dates:**

- Tuesday Oct. 10 - No Class (October Break)
- Thursday Nov. 23 - No Class (Thanksgiving Vacation)
- Midterm Exam I: Oct. 3, 8-10PM, LWSN B151.
- Midterm Exam II: Nov. 21, 8-10PM, LWSN B151.