STAT 512: Applied Regression Analysis
Spring 2013
TR   1:30 pm – 2:450 pm, UNIV 117

Instructor: Dr. Leonore Findsen
Office: MATH 535
Phone: (765) 49-46030 (Departmental office, messages only)
E-mail: LFindsen@Purdue.edu (or lfindsen@stat.purdue.edu)
Mailbox: in room MA 533 (open 8 am – 5 pm M-F)
Office Hours: TWR: 3:00 – 4:00

Textbook: Special reprint *of Applied Linear Statistical Models* (see bookstores) by Kutner, Neter, Nachtsheim & Li (KNNL). *(Required).* You may use the edition from this semester or from last semester which is slightly different. OR if you can find it *Applied Linear Statistical Models*, 5th *edition.* The text is large (and heavy with the special edition in 2 volumes, sorry!) and quite wordy, but it does provide alot of examples and graphs which are helpful. I can help you figure out which parts are really important and which parts you can skim over.

Applied Statistics and the SAS Programming Language, 5th edition, by Cody and Smith *(Optional).* A lighter book, this book provides a fairly readable introduction to the SAS Programming Language which could prove to be helpful during all parts of the course. However, most of the material in this book can be obtained online.

Web page: http://www.stat.purdue.edu/~lfindsen/stat512/stat512_sp13.html
   This page will be used to provide you with information relevant to the course, e..g. announcements, power point slides, homework assignments and solutions, data sets, dates of exams, review sheets, and changes to office hours. Please check this page regularly for updates. All announcements will be posted on the web page whether they are stated in class or not.

- **Course Description**: Thorough applied course in regression and analysis of variance including experience with the SAS statistical software package. Topics include inference in simple and multiple linear regression, residual analysis, transformations, polynomial regression, model building with real data, nonlinear regression. One-way and two-way analysis of variance, multiple comparisons, fixed and random factors, analysis of covariance. Not mathematically advanced, but covers a large volume of material. Requires calculus, and simple matrix algebra is helpful. Recommended for graduate students and for hard working undergraduates from all areas.

- **Course Goals:**
  1) Use the statistical methods of linear regression and ANOVA (ANCOVA) to analyze data.
  2) Perform diagnostics on the methods and perform appropriate remedial measures.
  3) Draw conclusions from these statistical analyses.
  4) Write statistical reports using correct terminology, analyses, and graphs.

- **Discussion Groups:** We will be using Piazza this semester for online discussions and questions. Besides other features, this allows both sections to interact in the same discussion group. This is the first time that I have used this site so please let me know if I need to change the setup, etc.. To enroll, please go to piazza.com/purdue/spring2013/stat512. You may post anonymously if you want; however, I will be giving bonus points for people who do post using their name. In addition, you can find classmates to form groups with which to study and for your project. (see below).

- **Blackboard**: Grades will be posted on Blackboard Learn. If there are any discrepancies of the grade on Blackboard and what you received on your paper, please contact me so that it can be changed. There will be no changes made to the grades on Blackboard after dead week (except for the final exam and material turned in during dead week).

- **Lectures**: PowerPoint slides will be used that contain some of the examples and figures used in lecture and will be provided in the Class Notes link on the class web site. No other class notes are supplied. Also, the sections of the chapters that I will be covering will also be on this page. We will cover roughly two chapters per week. You are always welcome to ask questions if I need to slow down.

- **Final Grade**: Homework (20%), Group Project (10%), Midterm 1 (20%), Midterm 2 (20%), Final (30%). The final percentages needed for a particular grade are as follows: 90 – 100 = A, 80 – 89 = B, 70 – 79 = C, 55 – 69 = D, < 54 = F. The minimum score needed for a given letter grade could be lowered if necessary but will not be raised. +/- grades are only given in special circumstances. (I expect – but won't guarantee – that the median grade in this class will be a high B.)

- **SAS Computer Software:** We will use SAS 9 to perform data analysis in this class. The intent of using software is to allow the computer to perform routine calculations and graphing, while we focus on choosing the appropriate analysis tools and interpreting the results. Computer software is NOT a substitute for understanding the statistical methods, and you will not have access to a computer during exams. SAS is available in the Purdue computing labs. You may also obtain a copy of SAS for your own PC for class purposes free of charge by showing your student ID in the lobby of the fifth floor of Young Hall (phone: 494-5100). Learning SAS will be one of the biggest challenges in this course, and you should be prepared to devote some time to this, especially in the first few weeks. The only way to learn how to use SAS is to try it! There are several sources of SAS help available. Make use of the SAS help system within the program to look up specific details. Another tool that should not be overlooked is the Web for finding SAS help. If you need help in person, the Statistics

Department provides a software consultant in MATH G-175, M-F, 10am-4pm (494-5100) (www.stat.purdue.edu/scs/). In addition, there are SAS evening help sessions: Wednesday TBA (www.stat.purdue.edu/academic_programs/courses/evening_computer_labs.html). There are also links on the SAS files Web page. In addition, you can get help from the instructor in office hours or by asking questions on Piazza. SAS manuals (besides the Recommended Text) are available in the bookstore. If you are looking for a easy to read book look for the optional books for STAT 501 (*SAS Essentials* by Elliott and Woodward).

- **Reading**: I expect you to read the text as we cover the material, which is about two chapters every week. It can help to read about a topic *before* it is covered in class. This does not mean that I expect you to learn it all on your own. Rather, your reading before the class should be a "first pass" at the subject. The first time through, I just want you to read through it quickly, in order to get a general idea of the material – the "big picture". Don't get bogged down in formulas or details; just try to get a rough idea of the material and get familiar with the vocabulary. This will prepare you for what is to come in the class, and will make the class easier to follow. If, as you are reading, you find something hard to understand, don't be alarmed or discouraged. Just make a note of any parts you found confusing, or any questions that occur to you as you read. Often, you will find that those questions are cleared up in the following class. If not, please ask during class! Later, as you are working on problem sets and studying for tests, you will find it helpful to read the material again. This time, read at a much more detailed level. It will be a lot easier to follow then, since you have already covered the material in class. Repetition and practice are important learning tools. The approximate sections that we will be covering and when are given in the Class Notes web page.

- **Exams**: There will be two midterm exams and the final examination. The two midterms will be in class. Midterm 1 is tentatively scheduled to be Thursday February 21 and Midterm 2 is tentatively scheduled to be on Thursday March 28. The time/date/location of the final will be available later in the semester. The Final Exam covers all of the material covered in the class with more emphasis on the material covered after the second Midterm. Each examination will contain both mathematical and conceptual (written or short answer) components. Each Midterm will be worth **20%** of your total grade and the final is worth **30%** of your total grade.

    Contact me as soon as possible if you are unable to take the exam at the scheduled time. After the exam key has been posted on the web site, NO MAKEUPS will be allowed.

- **Homework**: Homework will be due most weeks on Thursday during class. Please see the class web site for the dates and problem assignments as these will NOT be provided in class. The homework will normally be assigned at least one week in advance. You are encouraged to discuss the assignments with other students but you must write up your homework independently; identical solutions are NOT acceptable. Your homework must reflect YOUR understanding of the material. If identical solutions are found, all the parties involved will receive a 0 on that assignment. LATE HOMEWORK WILL NOT BE ACCEPTED UNDER ANY CIRCUMSTANCES; HOWEVER, THE LOWEST HOMEWORK GRADE WILL BE DROPPED. The key for each homework assignment will be posted on the web site shortly after it is due. Your total homework grade will constitute **20%** of your total grade in this class.

  Please do not pad your homework with endless printouts of SAS output. Only hand in those parts of the output that are directly relevant to your solution. You should edit any SAS output you plan to hand in by pasting it into an editor such as MS Word, and getting rid of extra space or unnecessary output. It is helpful to circle or highlight the portions of the output to which you refer in your solution. As a rule of thumb, only hand in what you actually expect the grader to read. SAS output should be pasted into your solution as you are answering the questions. Your SAS input file should be attached at the end of the homework. The input is not given a grade *per se*, but it can be helpful to the grader in trying to figure out what you did wrong and in assigning partial credit.

  Homework must be stapled if it is longer than one page. The first page must include your name, my name (Dr. Findsen), the number of the problem set (e.g., HW #2), the due date, and the course: STAT 512.

- **Project**: There will be a one group project due at the end of the semester in class on Thursday April 18. It will be helpful to form the groups of 2 – 4 people early in the semester to start to discuss what is necessary on the project as soon as possible. Please have one person of the group contact me (via e-mail) before the due date if an emergency will prevent you from turning in the paper on time. A late penalty of up to 50% may be applied. After the key for the project is posted, no late projects will be allowed. It might help to think of it as a directed final project done with limited guidance. Further details will be forthcoming and posted on the Homework web page. Your project grade will constitute **10%** of your total grade in this class.

- **Re-grades**: Since all humans make mistakes (including the professor and the grader), errors will occur in the grading. The following procedure is required if you want your assignment re-graded. Note: Only Type I errors should be submitted for re-grading, though we will accept Type II errors.

1) Attach a new piece of paper to the *front* of the work to be re-graded which contains the following information:
   a) The word "Re-grade" displayed prominently.
   b) Your name
   c) STAT 512. Be sure that this is clear so we know which course you are enrolled in.
   d) Which homework, project, or midterm is involved (e.g., HW 1)
   e) The relevant problem number(s) (e.g. problem 2a) or "Addition error"
   f) A detailed explanation of the suspected error ("Please look at problem 4" is **<u>NOT</u>** considered a detailed explanation) or provide the total number of points which you calculated if the assignment was added incorrectly.
   g) date of resubmission
2) Print out the solutions from the web page (at least the relevant portions) and circle the relevant piece of the solution. Attach this *behind* the work to be re-graded. (This is not required if you think that the points were added incorrectly.)
3) Give this packet to me, or put it in my mailbox (in MATH 533). A verbal explanation is neither necessary nor appropriate since a) I won't remember it, and b) the grader will do the re-grading anyway.

No exceptions will be made to this policy. The grader will be responsible for the re-grading (except for the midterms) and you will receive a written note from the grader explaining the outcome. I will review the grader's response before returning it to you to make sure the problem was resolved. Re-grade requests *must be submitted within 2 weeks of when the assignment is turned back in class unless explicitly stated otherwise by the professor.* If the above procedure is not followed, the re-grade request will be denied. Any rudeness accompanying a re-grade request will result in the assessment of a "technical foul" penalty equal to the total number of points for the disputed question. Note this is only for re-grading, if you are confused about the answer to a problem, please come to office hours (or send an e-mail) and ask.

- **General Comments:** This is not a math course. However, as in many other science courses, we will make use of mathematics quite extensively, and most questions will have some quantitative component. The use of SAS software will simplify many of the more computational tasks. However, the primary focus of this course is on learning how to do good science. Doing science well requires, among other things, a good experimental design and a correct and appropriate statistical analysis of the scientific data. Therefore, knowing when and when not to use a certain statistical method, and why, and how to interpret the results, are all at least as important as knowing how to actually carry out the calculations. In order to do well in this course, you must be prepared to master all of these areas.

- **General Course Policies:**
   1) If you have questions concerning the class; please come to my office hours, make an appointment, send me an e-mail or post on Piazza. I normally look at my e-mail numerous times during the day and evening hours and try to response promptly.

2) At the beginning of class, I normally make announcements, comments and provide advice. If you do need to arrive late or leave early, please sit in a location that does not disrupt the class.
3) The use of cell phones is prohibited in class and during the exams.
4) I strongly encourage all students to read the relevant material in class before attending the lecture so you are familiar with what will be covered. I also strongly recommend that you print out the relevant PowerPoint slides and bring them to class to take notes on.

- **Academic Dishonesty:** We take academic integrity very seriously in this course. The only true way to get an education is through hard work and striving to understand the concepts on your own. Penalties for academic misconduct range from a 0 on the assignment to failure in the course with referral to the Dean of Students for further sanctions. *Note that we punish not only the person who cheats but also the person who enables the cheater.*

- **Attendance:** Though I do not take roll in class, students are responsible for all material covered in the lectures. Not all of the material discussed in class is available in the textbook. If a student misses a class it is up to that student to check with their classmates to obtain the missed material.

- **Grief Absence Policy for Students:** Purdue University recognizes that a time of bereavement is very difficult for a student. The University therefore provides the following rights to students facing the loss of a family member through the Grief Absence Policy for Students (GAPS). GAPS Policy: Students will be excused for funeral leave and given the opportunity to earn equivalent credit and to demonstrate evidence of meeting the learning outcomes for missed assignments or assessments in the event of the death of a member of the student's family.

- **Violent Behavior Policy:** Purdue University is committed to providing a safe and secure campus environment for members of the university community. Purdue strives to create an educational environment for students and a work environment for employees that promote educational and career goals. Violent Behavior impedes such goals. Therefore, Violent Behavior is prohibited in or on any University Facility or while participating in any university activity.

- **Students with Disabilities:** Purdue University is required to respond to the needs of the students with disabilities as outlined in both the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990 through the provision of auxiliary aids and services that allow a student with a disability to fully access and participate in the programs, services, and activities at Purdue University.

It is the student's responsibility to notify the Disability Resource Center of an impairment/condition that may require accommodations and/or classroom modifications. The student may use the back of the questionnaire to inform the instructor at the beginning of the semester of any accommodations that are required; however, official paperwork is required at least two weeks in advance of any exams or assignments that require accommodations (except for the first homework assignment).

- **Health and other emergencies:** If a student has a serious medical issue, the instructor needs to be contacted via e-mail as soon as possible so accommodations can be made. If e-mail is not possible, leave a message at the number provided in the syllabus. The student will also need to provide documentation of the issue. Only limited accommodations can be made if the instructor is first contacted AFTER the student turns in the assignment or takes the exam.

  In the event of a major campus emergency or other circumstances beyond the instructor's control, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar. Relevant changes to this course will be posted onto the course website, http://www.stat.purdue.edu/~lfindsen/stat512/stat512_sp13.html, or can be obtained by emailing the instructor at LFindsen@purdue.edu. You are expected to read your @purdue.edu email on a frequent basis.

- **Nondiscrimination:** Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life.

  Purdue University prohibits discrimination against any member of the University community on the basis of race, religion, color, sex, age, national origin or ancestry, marital status, parental status, sexual orientation, disability, or status as a veteran. The University will conduct its programs, services and activities consistent with applicable federal, state and local laws, regulations and orders and in conformance with the procedures and limitations as set forth in Executive Memorandum No. D-1, which provides specific contractual rights and remedies.

- **Approximate Outline** (including the approximate readings for each section. This information will also be located on the Class Notes web page which will contain the most up-to-date version.)

Topic #	Topic	Reading
1	Simple Linear Regression and Inferences	1.3, 1.6, 1.7, 2.1 – 2.10
2	Regression Diagnostics, Remedial Measures, Misc. Topics	3.1 – 3.3, 3.8, 3.9, 4.1 – 4.6
3	Matrices, Simple and Multiple Regression	5.1 – 5.7 (review), 5.8 – 5.12, 6.1 – 6.9
4	Multiple Regression, Qualitative Predictors, Model Selection	7.1 – 7.4, 7.6, 8.1 – 8.3, 8.5, 9.1 – 9.4
5	Diagnostics and Remedial Measures	10.1 – 10.5, 11.1, 11.2
6	One-way ANOVA	16.1 – 16.8, 17.1 – 17.7, 17.9, 18.1 – 18.6
7	Two-way ANOVA	19.1 – 19.10, 20.1, 20.2
8	Two-way ANOVA, Three-way ANOVA	23.1 – 23.3, 24.1, 24.3 – 24.5
9	Random vs. Fixed Effects	25.1 – 25.3
10	One-way Analysis of Covariance	22.1 – 22.4
11	Nonlinear Regression, Logistic Regression, Generalized Linear Models	13.1, 13.2, 13.5, Selected parts from: 14.1 – 14.5, 14.8, 14.10 -14.11, 14.13 - 14.14

This syllabus is subject to change.