



Mathematical Sciences 807
Applied Multivariate Analysis
Instructor: Calvin L. Williams, Ph.D.
Fall 2011 MWF 1:25-2:15 p.m. Martin M-102

Office: 0-323 Martin Hall

Telephone: (864)-656-5241

E-mail: calvinw@ces.clemson.edu

Office Hours: MTWTH 9:30-10:30 or By Appt.

Course Web Page: <http://www.ces.clemson.edu/~calvinw/mthsc807.html>

I. Required Materials:

- (i) **Text:** Methods of Multivariate Analysis - Alvin Rencher
- (ii) *Supplementary:* Multivariate Statistics with R- Paul Hewson (Highly Recommended!)
- (iii) *Supplementary:* Online SAS manuals for programming in SAS.
- (iv) *Supplementary:* Applied Multivariate Statistics with SAS Software-Authors: Khatree and Naik

II. Course Description: Multivariate data through experimentation and observation occur quite often in engineering, business, social sciences, as well as biological and physical sciences. This is a course in applied multivariate data analysis. It will cover descriptive and graphical methods for continuous multivariate data, the multivariate normal, multivariate tests of means, covariances and equality of distributions, univariate and multivariate regression and their comparisons, multivariate analysis of variance, covariance structure models, and discrimination and classification. Furthermore it should be emphasized that this course and hence the chosen text, is designed around the application of multivariate techniques to continuous data, time allowing we will endeavor to discuss methods of discrete multivariate analysis from prepared class notes. Students will learn how to use statistical software to facilitate the understanding of the foundations of multivariate analysis. Statistical packages will include R, MatLab, and SAS.

III. Prerequisites: Students taking MthSc 807 are expected to have had a course in regression or intermediate data analysis (perhaps 805) and exposure to some statistical software although it is not required. The equivalence of MthSc 403/603 (Intro to Statistical Theory) can be considered preparatory for those students interested in multivariate data analysis. Prerequisites are a working knowledge of general linear models, statistical inference concerning these types of models, and hypothesis testing, and elementary matrix operations.

IV. Attendance Policy: All classes should be attended, but, if you are ill stay at home. I will accept e-mail or phone messages to that effect. Note that this does not exempt you from turning in homework/projects on time nor taking quizzes at their proposed times. Legitimate excuses must be offered with respect to the day(s) missed. Attendance will be monitored. It is to the instructors discretion whether an excuse is legitimate or not. Accordingly, the university's policy on religious holidays will be acknowledged and honored.

V. Tardy Professor Policy: If the instructor is more than 15 minutes late for any class you may leave.

VI. **Grading Policy:** The two regular exams (midterms if you wish) and a final exam which will count as 50% of the final grade, homework sets 10%, several multivariate exploratory data cases 20%, and a final multivariate data analysis project 20%. The final exam will cover the more important topics covered during the semester and will consist of several short answer questions requiring no more than a calculator.

VII. **Grading Scale:**

A 100 - 90; B 89 - 80; C 79 - 70; D 69 - 60; F 59 - 0.

VIII. **Examination Policy:**

There will be **two** 50 minutes in class **closed book** examinations and a **closed book final** examination consisting of short answer questions requiring no more than a calculator. **No makeup examinations will be given.** Any student who misses an examination without a **legitimate excuse**, *ie, a documented medical excuse*, will receive a score of **zero** for that exam. A student with a **legitimate excuse**, will receive a final score based on all other class work. More than one missed exam will require withdrawal from the course and/or the receipt of a failing final grade.

IX. **Homework and/or Take Home Projects:**

There will also be several homework sets and/or take home projects assigned from the text as well as from material covered during class. It is imperative that each student be completely comfortable with these assigned problems and projects. Homework is a critical part of the course. It will include both theoretical and applied work. Late homework will not be accepted (although you may turn in assignments early). Homework should be detailed enough to adequately demonstrate your solution. You may discuss the homework problems with other students, however, the final work you turn in must be your own. **Don't copy other students' solutions.** Questions regarding them can be asked during office hours and via email. **All homework assignments must be sent electronically. This requires that solutions to homework sets and take home projects be typed. No hand written assignments will be accepted. All data sets will be available electronically to facilitate this being done. Homework sets will be due two class periods after they have been assigned.**

X. **Final Multivariate Data Analysis project**

The purpose of this project is to have the student conduct a complete and thorough analysis of a multivariate data set. Data sets can be selected from journal articles or other referenced sources. The analysis should be completely novel from the approach taken in the source.

The final project should include:

1. A one page description of the data set, which should contain at least 50 observations and 5 variables, including the source.
2. A one page proposed analysis, methods, and rational for the methods used.
3. The results of the analysis, including a printouts and graphics. Graphics should be incorporated into results. Source code should be included in the appendices along with a reference page sourcing the problem and the data.
4. Data from the text book will not be allowed, although simulation of data similar to those in the text will be allowed.

Reports should be neatly typed, well-organized and attractive. Graphical displays (either computer-generated or hand-drawn) are encouraged. Generally, graphs are more effective if they are incorporated into the text, rather than hidden at the end of the report. You may also use a computer package to aid in the data analysis. If you do so, the results should be discussed in the text of your report, and the computer output itself may be included in an appendix.

A typed rough draft of the final report will be due approximately 2 weeks before the final report is due.

The project is worth 100 points. Grades will be based on:

- Appropriate and correct procedures 50 pts;
- Well-written and attractive presentation 20 pts;
- Grammar, spelling and punctuation 20 pts;
- Complexity 10 pts.

XI. **Cell phones should be unseen and unheard during class.** You **are not** the exception to this rule. Students who fail to comply will be asked to leave the classroom. Please turn your phone off and store it away when you enter the classroom. It should not be left on your desk during class.

XII. **Academic Integrity:** The following is the official statement on “Academic Integrity.” The Provost urged it’s placement in the syllabus, cited in the Graduate Announcements, page 27. “As members of the Clemson University community, we have inherited Thomas Green Clemson’s vision of this institution as a “high seminary of learning.” Fundamental to this vision is a mutual commitment to truthfulness, honor, and responsibility, without which we cannot earn the trust and respect of others. Furthermore, we recognize that academic dishonesty detracts from the value of a Clemson degree. Therefore, we shall not tolerate lying, cheating, or stealing in any form.” “When in the opinion of a faculty member, there is evidence that a student has committed an act of academic dishonesty, the faculty member shall make a formal written charge of academic dishonesty including a description of the misconduct, to the Dean of the Graduate School. At the same time, the faculty member may, but is not required to, inform each involved student privately of the nature of the alleged charge.”

XIII. **Disability Access Statement:** Students with disabilities who need accommodations should make an appointment with me to discuss specific needs within the first two weeks of classes. Students should present a Faculty Accommodation Letter(FAL) from Student Disabilities Services when we meet. Student Disability Services is located in G-20 Redfern (www.clemson.edu/sds/). Please be aware that accommodations are not retroactive and new FAL must be presented each semester.

Lecture Topics(tentative)

Aug 24	3.1-3.4	Introduction and Characterization of Multivariate Data
Aug 26	3.5-12	Characterization of Multivariate Data
Aug 28	3.5-12	Characterization of Multivariate Data
Aug 31	3.5-12	Characterization of Multivariate Data
Sep 2	4.1-4.3	Multivariate Normal Distribution
Sep 5	4.4-4.5	Multivariate Normal Distribution
Sep 7	4.4-4.5	Multivariate Normal Distribution
Sep 9	5.1-5.3	Test on Mean Vectors
Sep 12	5.4-5.6	Test on Mean Vectors
Sep 14	5.7-5.8	Test on Mean Vectors
Sep 16	5.9-6.2	Test on Mean Vectors
Sep 19	5.9-6.2	Test on Mean Vectors
Sep 21	6.3-6.4	Multivariate Analysis of Variance
Sep 23	6.5-6.7	Two-Way Classification
Sep 26	6.5-6.7	Two-Way Classification
Sep 28	6.8-6.11	Profile Analysis
Sep 30	6.8-6.11	Growth Curves
Oct 3	7.1-7.4	Tests on Covariance Matrices
Oct 5	7.1-7.4	Tests on Covariance Matrices
Oct 7	8.1-8.5	Discriminant Analysis
Oct 10	8.1-8.5	Discriminant Analysis
Oct 12	8.6-8.9	Discriminant Analysis
Oct 14	8.6-8.9	Discriminant Analysis
Oct 17	Fall Break	
Oct 19	9.1-9.4	Classification Analysis
Oct 21	9.1-9.4	Classification Analysis
Oct 24	9.1-9.4	Classification Analysis
Oct 26	10.1-10.4	Multivariate Regression
Oct 28	10.1-10.4	Multivariate Regression
Oct 31	10.1-10.4	Multivariate Regression
Nov 2	10.1-10.4	Multivariate Regression
Nov 4	10.1-10.4	Multivariate Regression
Nov 7	12.1-12.2	Principal Component Analysis
Nov 9	12.3-12.5	Principal Component Analysis
Nov 11	12.6-12.8	Principal Component Analysis
Nov 14	12.9	Principal Component Analysis
Nov 16	13.1-13.2	Factor Analysis
Nov 18	13.3	Factor Analysis
Nov 21	13.4	Factor Analysis
Nov 23	Thanksgiving	
Nov 25	Thanksgiving	
Nov 28	13.5	Factor Analysis
Nov 30	13.6-8	Factor Analysis
Dec 2	14.1-14.2	Cluster Analysis
Dec 5	14.3	Cluster Analysis
Dec 7	14.4	Cluster Analysis
Dec 9	14.5	Cluster Analysis