

Office Hours: M&W 2:00-4:00
375H Melcher, 713 743-4175
Home phone 713 663-7012
JHESS@UH.EDU

Course Objective:

To learn multivariate statistical methods that uncover surprising but valid linkages between variables and explain and predict their measured values.

This should support your personal objective as a doctoral student, “to learn to do research that is publishable in the top academic journals in your discipline.” Specifically, you will learn

- to match multivariate methods with academic research.
- to perform multivariate data analysis (using SPSS) including:
 - Multivariate Analysis of Variance
 - Discriminant Analysis
 - Multivariate Regression Analysis
 - Factor Analysis
 - Multidimensional Scaling
 - Logit Models
 - Structural Equation Modeling
 - Cluster Analysis
- to interpret the results and test the assumptions of a multivariate data analysis.
- to understand academic research employing multivariate techniques.

The course is introductory in nature but aimed at moving Ph.D. students toward the frontier of empirical techniques as used in academic journals. Confidence and proficiency do not come with the first step on this journey. The goal of this course is to set you out on a path that will allow you to continually acquire additional expertise. It’s a long journey. In 2025, you’ll likely turn back to the textbook that you first read in 2005, and still learn something valuable.

This “journey” metaphor brings to mind several things. You are journeying into a land where a new language is spoken: algebraic manipulation of data stored in matrices. To be a skilled researcher, you must become conversant, if not fluent, in this matrix language. Some students in this course have already had exposure to **matrix algebra** in undergraduate math courses, but others will be seeing it for the first time. Both types will succeed in this course, but for those who find matrices strange and unusual, let me encourage you to take extra effort to master these algebraic techniques. It will pay great dividends in the years to come, when you feel confident to attack research that uses new data analytic methods. For this reason, an introductory problem-oriented book on matrix algebra is listed below.

You are bound to get lost occasionally in the land of multivariate statistics. I did. I still do. However, you have the opportunity to take advantage of the close personal relationships between faculty and doctoral students here at Houston. I personally offer any help you might need at any point in your development as a researcher and hope that we develop close connections this semester that will stand the test of time, so you may always come back to me with questions about how to find our way in a complicated research terrain. Do not be shy!

“The most beautiful thing we can experience is the mysterious.” Albert Einstein

Notice the word “surprising” in the above course objective. To become well-known researchers, we must continually seek mysterious and surprising results. However, science is a skeptical discipline. Many a scary “cold fusion” story has been told round the academic campfire (http://en.wikipedia.org/wiki/Cold_fusion#History). You must balance the desire to discover unforeseen relationships with the requirement that these be valid and reliable findings. Multivariate statistics helps you do this.

Learning Strategy

Just do it!

Nike advertisement

Problem Sets

My learning philosophy is expressed by the Chinese proverb, “I hear; I forget. I see; I forget. I do; I remember.” To help you develop familiarity with the multivariate data analysis, almost daily you will do homework drawn from the end-of-chapter exercises. These involve real data. Much of class will be spent understanding the proper answers to these.

You must learn how to use computers to calculate the statistics that are most appropriate for a problem. The more proficient you are using a statistical package the more effective you will be as a research scholar. We are going to use SPSS throughout this course, and one of the course goals is to help you become expert at doing multivariate statistics in SPSS. Almost every class meeting I will be doing data analysis with SPSS to repeatedly demonstrate that the abstract analysis of the textbook becomes real when data is confronted. If you have a laptop computer with SPSS, then you too can get dirty hands from digging in the data.

The problem sets can sometimes be very challenging. To promote efficient learning, follow “**Hess’s Rule**”: if you have not been able to answer a problem after working on it for three hours, you must call me - at the office, at home, in the morning, or late at night - to get a hint or two. Seriously! My office and home phone numbers are listed above, but I am also in the telephone directory.

How good are your skills at typing mathematical and statistical equations? To enhance them, I’d like you to write-up your homework exercises answers using MS Word, using the Equation Editor to express mathematical formulas. If you have no experience with this, there is a nice tutorial at <http://www.ist.uwaterloo.ca/ec/equations/equation.html> with videos. You can cut and paste output from SPSS into your Word document.

Term Project

If you are going to travel in strange land where a new language is spoken, there has to be strong motivation for bearing the cost. The obvious motivation is that you want to do research just like the intellectual leaders of your discipline do, and many of them use these multivariate techniques. You will be most motivated if you are not studying these multivariate methods in the abstract, but studying them in the actual practice of your discipline.

To that effect, **by September 7 please find an empirical article that you admire written by one of your professors and from her/him get the multivariate dataset.** If you are having problems locating a paper and dataset, please come see me, and we can discuss alternatives. You will analyze this data using the methods described in this course and make a presentation to the class at the end of the term. You may want to replicate the published analysis or to use another multivariate method to extend the study. In either case, you are to estimate the model, assess overall fit, and draw conclusions from the findings. You will get to communicate your analysis to the class as you would at an academic conference (15 minute presentation followed by 5 minutes Q&A) and in a paper (10 page maximum, double-spaced, excluding figures and tables).

I’d like to meet you individually for 10-15 minutes on **September 14** to discuss what you intend to do with this dataset (time schedule to be determined later). Please bring a clean photocopy of the paper for me at this meeting, but also send me an e-mail message with attachment of the dataset and with an electronic copy of the paper (almost all journals are now available electronically on the library web page but if not, the author may have a computer file with the paper).

Course Grading

Like in all doctoral courses, our goal is for you to do research that is publishable in the top academic journals (whose editorial boards are also receiving submissions from the best “veteran” researchers in the world). If your statistical analysis looks inexpert, how will you convince the editor that you have truly uncovered surprising but valid linkages between variables? I really want you to succeed in this endeavor, so I will provide evaluations of problem sets and project from the perspective of a journal editor. Personal improvement is all that matters to you and to me, but because the university requires that I give you a grade from a standardized set of options, I will combine my evaluations using the following weights: homework problems 50%, project presentation 10%, and project paper 40%.

Textbooks

Required: Analyzing Multivariate Data by James Lattin, Douglas Carroll and Paul Green, 2003, ISBN 0-534-34974-9, Thomson Learning.

Recommended: SPSS Companion for Lattin/Green/Carroll's Analyzing Multivariate Data, ISBN0-534-38226-6

Recommended (for those without background in matrix algebra): Matrix Operations, Richard Bronson, Schaum Outline Series, McGraw-Hill, 1989.

Comments: If matrix algebra is new to you, the Schaum outline book has hundreds of worked out examples of matrix methods (and it is cheap). I have not asked the bookstore to order the Schaum or SPSS Companion books, but you can buy them from web bookstores.

SPSS Statistical Software

- Computer applications will use **SPSS 13.0** for Windows. Later in the semester, we will use **AMOS** to do structural equation modeling. Your department may provide these for you. SPSS has bundled SPSS 13.0, Regression, Advanced Models, and AMOS into a Grad Pack and at <http://www.journeyed.com/itemDetail.asp?T1=62581573+R> a student may buy a 4 year license for this for \$199.98. If you do not want to have a 4 year license, a one year license for SPSS 13.0 is available for \$80 at UH Discount Software for “staff” (<http://www.uh.edu/infotech/php/home.php>). However, AMOS for one year will cost you an additional \$80. Central Site Computing Center located in room 110 of the Social Work building has SPSS on some computers. Hopefully we can get it on computers in RICS labs.
- My “simpleton’s guides” will be posted on our WebCT conference (see below). As the semester progresses, I will produce snapshot tours of SPSS for specific data analytic techniques; that is, I’ll make a screen capture after each step of using SPSS and post a folder of these JPEG files that you can view as a slideshow.
- If you go to Central Michigan’s website <http://www.cst.cmich.edu/users/lee1c/spss/> you can view and listen to nice tutorials for many SPSS methods. It will require a Quicktime Movie player, but that can be downloaded free (see instructions on the tutorial web page). Other such websites are <http://www.utexas.edu/its/rc/software/> and <http://www.spsstools.net/>.
- If you would like to see innovative statistical research software explained, register at <http://www.spss.com/downloads/> to download SPSS’s white papers on a variety of topics.

WebCT

We will use the WebCT system as a bulletin board to facilitate electronic communication. On our MARK 8349, I will post datasets, SPSS simpleton’s guides, lecture notes for some topics, **homework exercises**, and respond to your questions about statistical research. You can log onto WebCT from any computer that has Web access to <http://uh.edu/webct/index.html>.

Schedule of Topics and Readings

Below is a tentative schedule of topics. I may decide to reduce the number of topics in order to cover the important topics less hurriedly, depending on how easy/difficult you find the material.

Schedule				Topic	Readings
1	M	Aug	22	Love the data... Be the data - SPSS beginner's tutorial: descriptive stats, explore, recoding, computing, splitting dataset, scatterplots with rotations, boxplots, outliers, missing data	1 (chapter in text) SPSS notes on WebCT
2	W	Aug	24	Matrix manipulation of data	2, Notes on Matrix Algebra on WebCT
3	M	Aug	29	SPSS: matrix computations	SPSS Notes on WebCT
4	W	Aug	31	Multivariate normal random variables	Notes on WebCT
	M	Sept	5	Labor Day - no class	
5	W	Sept	7	Random samples from multivariate normal	Notes on WebCT
6	M	Sept	12	Inferences about a mean vector	Notes on WebCT
7	W	Sept	14	Classical Linear Regression Model: Gauss-Markov	3.1-3.3
8	M	Sept	19	Violations of classical regression assumptions	3.4
9	W	Sept	21	To be rescheduled - no class	
10	M	Sept	26	Regression topics: multicollinearity, non-spherical error, dummy variables, mediator variables, Poisson regression	Notes on WebCT
11	W	Sept	28	Specification errors in regression	Reading from Lovell
12	M	Oct	3	Exploratory Factor analysis	(skim 4), 5
13	W	Oct	5	SPSS applications of factor analysis	SPSS Notes on WebCT
14	M	Oct	10	Multidimensional Scaling	7
15	W	Oct	12	Cluster analysis	8
16	M	Oct	17	Confirmatory Factor Analysis	6
17	W	Oct	19	Confirmatory Factor Analysis via AMOS	AMOS notes on WebCT
18	M	Oct	24	Simultaneous equation regression and 2SLS	Notes on WebCT
19	W	Oct	26	Structural equation models (SEM)	10
20	M	Oct	31	Identification of systems of equations	Handout
21	W	Nov	2	Structural equation models via AMOS	AMOS notes on WebCT
22	M	Nov	7	ANOVA and MANOVA	11
23	W	Nov	9	MANOVA via SPSS General linear model	Notes on WebCT
24	M	Nov	14	Discriminant analysis	12
25	W	Nov	16	Binary logit models of qualitative choice	13.1-13.3
26	M	Nov	21	Multinomial Logit	13.4-13.5
	W	Nov	23	Thanksgiving Vacation – no class	
27	M	Nov	28	Presentation of project data analysis 1	
28	W	Nov	30	Presentation of project data analysis 2	

Academic Honesty: The University of Houston Academic Honesty Policy is strictly enforced by the C. T. Bauer College of Business. No violations of this policy will be tolerated in this course. A discussion of the policy is included in the University of Houston Student Handbook, <http://www.uh.edu/dos/hdbk/acad/achonpol.html>. Students are expected to be familiar with this policy.

Accommodations for Students with Disabilities: The C. T. Bauer College of Business would like to help students who have disabilities achieve their highest potential. To this end, in order to receive academic accommodations, students must register with the Center for Students with Disabilities (CSD) (telephone 713-743-5400), and present approved accommodation documentation to their instructors in a timely manner.