PSYC 802: Multivariate Statistics
Queen's University - Winter 2007
Humphrey 223
Mondays 11:30 – 1:00
Wednesdays 1:00 – 2:30
Lab: Fridays 8:30 – 10:00 Humphrey 219

Instructor: Dr. Tom Hollenstein (tom.hollenstein@queensu.ca) Craine 220
Office Hours: Wednesdays 2:30 – 3:30 and by appointment
TA: Jamal Mansour (4jkm@qlink.queensu.ca) Humphrey 324
Office Hours: 12:30-1:30 Tuesdays

Textbook:
Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. New York: Pearson

Other Readings:
Abelson, R. P. (1995). *Statistics as Principled Argument*. Hilsdale, NJ: Erlbaum. pp. 54 – 103.
Cohen, J. (1990). Things I have learned so far. *American Psychologist, 45,* 1304-1312.
Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 49,* 997-1003.
Garson, G. D. (2006). Reliability Analysis. http://www2.chass.ncsu.edu/garson/pa765/reliab.htm.
Levesque, R. (2006). *SPSS Programming and Data Management, 3rd Edition: A Guide for SPSS and SAS Users*. Chicago: SPSS. **Chapters 2 – 7**.
Rosnow, R. L., & Rosenthal, R. (1989). Definition and interpretation of interaction effects. *Psychological Bulletin, 105,* 143-146.

These and other recommended readings can be found on Web CT course site.

***Course Description*.**

In this course, we will cover the **concepts**, **procedures**, and **interpretations** of several multivariate methods. I assume you already have a good grasp of univariate methods and issues so that we may delve into the issues that arise when you need to analyze two or more dependent and/or independent variables. After covering the basics of data cleaning, reliability, and the computational language of matrix algebra, we will cover each of the three major multivariate methods: factor analysis, MANOVA, and regression. These three are mathematically related to each other and most other techniques can be understood as variations of these three. Weekly labs will focus on SPSS procedures as well as clarify issues from lecture and the homeworks.

Although statistics are based on mathematical formulas that represent the relationships among variables, the intent of this course is to focus on statistics as a means of principled argument (Abelson, 1995). We use statistics to make inferences about the true nature of the world, to answer research questions, to test theories. Hence, the goals of the course are to make sure that you walk away understanding the **conceptual underpinnings** of each technique, the SPSS **procedures** necessary to conduct these analyses, and the skills to be able to critically **interpret** your own results and the claims of the research you encounter throughout your careers.

A few other things for your consideration:

1. The range of expertise in the class is broad. I will aim for the middle level. Thus, advanced students may be interested in more detail and novice students may struggle a bit. That is the nature of such a course as this. However, I have always felt like I could take one introductory stats course each year and still get something out of it. Which leads me to…
2. Redundancy. The absolutely best way to learn statistics is through redundancy, a repetition of the same ideas, multiple presentations of the same material, re-experiencing a technique in different contexts, the reiteration of crucial details, and practice, practice, practice. To this end, I will emphasize what is shared among the techniques as well as try to present the information in several ways, when possible (i.e., equations, graphs, examples). However,
3. I am not going to be able to impart to you absolutely everything about multivariate statistics in this course for several reasons: (a) not enough time – each technique could take 13 weeks on its own; (b) each research question and data set can present relatively unique issues for analysis; (c) some issues are still hotly debated among the statistical intelligentsia; and (d) I simply don't know absolutely everything (my omniscience is less than complete but I do enjoy the challenge). However, this is true of any course and instructor. What I *will* be able to give you are the fundamentals so that when you need to run a complicated analysis 8 years from now, you know where to look for a refresher, or be able to understand how to interpret the results, or at least know how to ask questions of a statistician. Thus, we will need to be able to speak a…
4. mathematical language. There are about 20 symbols and a few labeling conventions we require in order to efficiently communicate about and manipulate multivariate data. Yes, this really is necessary. If you think you are not mathematically inclined, then think about this as a language with very few words. I have chosen an exceptionally clear text and I will strive for that clarity in my lectures as well. The goal is to get you *thinking* "multivariately". Still, no matter what I do, some of you will be…
5. anxious. Statistics has a bad rap. As a result, many people approach it with fear and loathing. I have seen many students dig in their heels and put more effort into fretting, resisting, avoiding, or blaming than just trying to learn. If this describes you in any way, I implore you to suspend your apprehensions, breathe deeply, come see me, look for alternate sources of information, ask questions, and keep at it. I promise, multivariate statistics are good for you.

Course Requirements.

Registered students are expected to attend every class and lab. The readings come mostly from the text (Tabachnick & Fidell, 2007), but there will also be a few short articles required as well. Because most of you have not encountered any of the techniques in this course so far, we will have to maximize redundancy (see #2 above). Therefore, it is imperative that you read the required material *before* each class and probably reread it again afterwards.

There will be a total of 8 homeworks and 2 take home exams. The homework will be computational and the exams will be conceptual (i.e., tests of your understanding of

the central concepts and functional character of the techniques). The final exam will mostly cover the second half of the course, but some of the larger concepts from the first half of the course will also be fair game.

Homework. On Mondays, the TA will post the homework for that week on WebCT and it is due the following Monday in class. You will be able to discuss the assignment during lab on Wednesday afternoon. Typically, the assignment will come with a data set (also on WebCT) with several instructions for analysis. Your homework is complete if you run all the required analyses, edit the output files (I do not want all of the output, only the relevant tables and graphs), annotate (label) them, and write up the results. Please use Tabachnick & Fidell's result summaries as models. You must write clearly, concisely, and meaningfully. You must have a **paper copy** of the homework to the TA by the start of class on Monday (5 points off for every day late). More details about the homework procedures will be discussed during the first lab.

Homework 1: Syntax, Exploratory Data Analysis, Missing Data
Homework 2: Matrix Algebra
Homework 3: Data Reduction
Homework 4: Factor Analysis
Homework 5: MANOVA
Homework 6: Repeated Measures and mixed-model MANOVA
Homework 7: Hierarchical Regression
Homework 8: Integration: Multiple Regression, Logistic Regression, and DFA

Grading

Homeworks (8 x 20 points each) =	160
Take home midterm =	80
Take home final =	120
Total =	360 points

Electronic Considerations.

The course materials will be distributed through WebCT. Please log in asap to make sure that you have no problems with access. Those who are auditing the course will need to see Jane Dauncey in the Psychology main office – she will get you on the course site.

We will be using SPSS for all analyses in this course. Currently, Queen's supports version 16. I believe that versions 14 and 15 will perform all the analyses required. The lab in 219 has copies on each computer for you to use. However, I recommend that you have other access to SPSS if you don't already. I believe the license is $75 for one year and it is well worth it. Let me know if you have any problems.

EMAIL

If you have a question or problem, please email the TA *first*. In all email exchanges, we ask that you include "PSYC802" in the subject heading. We will create a class email list so that you may contact each other.

Course Schedule

Week	Day	Date	Topic	Reading*	Due
1	Monday	Jan. 7	Orientation		
	Wednesday	Jan. 9	The Basics & Overview	1 & 2	
	Friday	*Jan. 11*	*Lab: Orientation*		
2	Monday	Jan. 14	Data Management	4; Abelson	
	Wednesday	Jan. 16	Matrix Algebra I: intro	App. A	
	Friday	*Jan. 18*	*Lab: Missing Data*	*3; Lavesque*	
3	Monday	Jan. 21	Matrix Algebra II: stats		HW 1
	Wednesday	Jan. 23	Reliability Analysis	Garson	
	Friday	*Jan. 25*	*Lab: Matrices*		
4	Monday	Jan. 28	Principal Components	13	HW 2
	Wednesday	Jan. 30	Factor Analysis I	13	
	Friday	*Feb. 1*	*Lab: Data Reduction*		
5	Monday	Feb. 4	Factor Analysis II	13	HW 3
	Wednesday	Feb. 6	GLM (M)ANOVA intro	17; R&R; Cohen	
	Friday	*Feb. 8*	*Lab: ANCOVA & GLM*	*6*	
6	Monday	Feb. 11	MANOVA I: concepts	7	HW 4
	Wednesday	Feb. 13	MANOVA II: SPSS	7	
	Friday	*Feb. 15*	*Lab: MANOVA*		
7	**Feb. 18-22**		**READING WEEK**		
8	Monday	Feb. 25	MANOVA III: wrap up	7	Midterm
	Wednesday	Feb. 27	Discriminant Functions	9	
	Friday	*Feb. 29*	*Lab: DFA and MANOVA*		
9	Monday	Mar. 3	Repeated Measures I	8	
	Wednesday	Mar. 5	**NO CLASS**		
	Friday	*Mar. 7*	*Lab: Repeated-measures*		
10	Monday	Mar. 10	Repeated Measures II	8	HW 5
	Wednesday	Mar. 12	Canonical Correlation	12	
	Friday	*Mar. 14*	*Lab: Regression I*		
11	Monday	Mar. 17	Multiple Regression Intro	5	HW 6
	Wednesday	Mar. 19	Multiple Regression in SPSS	5	
	Friday	*Mar. 21*	*Good Friday – No Lab*		
12	Monday	Mar. 24	Multiple Regression	5	HW 7
	Wednesday	Mar. 26	Logistic Regression	10	
	Friday	*Mar. 28*	*Lab: Regression II*		
13	Monday	Mar. 31	Integration & Other Methods		HW 8
	Wednesday	April 2	Wrap Up		
	Friday	*April 4*	*Lab: Review*		
FINAL	***Wednesday***	***April 11***	***Take Home Final Due 5pm***		

*Numbers refer to chapters in Tabachnick and Fidell (2007)
Lavesque = SPPS guide to data management and programming (2006) Chapters 2-7
R&R = Rosnow & Rosenthal (1989)
Cohen = Cohen (1990, 1994)