CJ 742: Advanced Statistics I

Instructor: Professor Scott Menard
Office phone: 936-294-3280 (leave a message and wait a few days; or...)
Cell phone: 303-818-3165 (...the number you want to use for a quicker response...)
Fax: 936-294-1653 (...probably not the fastest or most reliable method; or...)
Email: ScottMenard@shsu.edu (most reliable means of communication with me)
Office: CJC C-215

Office hours: TuTh 1:00-1:50, We 1:00-2:45 PM, and Th 3:30-4:00 PM; and by appointment.

Class times: We 3:00-5:50 PM

Location: CJC A-205

Course Objectives:

(1) To prepare you for the comprehensive examination in statistics.
(2) To prepare you to do some real data analysis, using multivariate techniques appropriate for the level of measurement of the dependent and independent variables, including techniques for assessing reliability of measurement as well as techniques for assessing how well a set of independent variables predicts a dependent variable.
(3) To prepare you for more advanced courses in statistics.

Course Philosophy

Students in this course will typically be graduate students in criminal justice (possibly with backgrounds in other social or behavioral sciences) who have had a first course in statistics (frequency distributions, descriptive statistics, contingency tables or crosstabulation, bivariate correlation and regression, nonparametric statistics, and possibly some analysis of variance), and who are ready to learn the more advanced multivariate techniques that form the backbone of statistical analysis in social science research. There is a brief and selective review of basic statistics, focusing on concepts that will be emphasized in multivariate statistical analysis, but students are expected to know most or all of this material already.

This course emphasizes the application of statistics to the analysis of social science data. By the end of this course, you should understand
(1) how to select the appropriate multivariate statistics to analyze a research problem,
(2) how to compute the statistics using SPSS in a Windows environment, and
(3) how to interpret the statistical output from SPSS to answer research questions and test research hypotheses.

The mathematical content of the course is intended to serve these three goals; it is not the focus of the course. If you are interested in taking a course mathematical statistics (something I would recommend, especially if you intend to specialize in quantitative methods and statistics at

the graduate level), you really need to look into the Mathematics department. In this course, it is not only (or even primarily) the mathematics and computer applications, but also the substantive content and the ability to translate the computer output into social-scientifically meaningful English prose, that are important.

The Statistics Lab (CJ 615)

Periodically (about once every two weeks or so) we will meet in the Criminal Justice Statistics Lab in CJC A111 for sessions involving hands-on practice in using SPSS; and you will need access either to the lab or to a site licensed copy of SPSS in order to complete the computer assignments for the course. Based on past experience, I suspect that this is where most of the real learning will occur.

Grading and Assignments

(1) Computer Assignments and Examinations. There will be four computer assignments and four examinations, plus your choice of a final (fifth) computer assignment or a final (fifth) examination. The first computer assignment and examination are on the review material. The first four examinations will be given the same day each computer assignment is due. The final computer assignment is due the scheduled day of the final examination; if you really want to, you can do both, but you only are required to do *either* the fourth computer assignment *or* the final examination, unless you need to make up an examination (see below). Each of the examinations will be designed to last 30-60 minutes. Examinations will be open book, open notes, but you are not allowed to compare answers during the examination. The computer assignments and the examinations will cover roughly the same material.

After the first computer assignment, your familiarity with the computer will be assumed; it is up to you to take whatever measures are necessary to make this assumption correct. Appropriate measures include doing the reading, asking me questions, and seeking whatever outside help is necessary, including tutoring if appropriate. Computer assignments will be due two weeks after they are assigned - at which time you will get your next computer assignment.

You are welcome (actually, encouraged) to work together and help each other in learning how to run SPSS and how to interpret the results. The final presentation of the work, however, must be your own. If you have someone else run SPSS and write up the results for you, you will get nailed on the examinations.

The examinations will not require you to perform calculations. That is what computers are for. Instead, examinations will focus on (1) choosing the appropriate statistical procedures to analyze a problem, and (2) interpreting results of SPSS output for statistical procedures as applied to a specific problem. (I told you that if someone else did your computer assignments you would get nailed on the exams, remember?)

(2) Final Examination. The final examination is a comprehensive final examination, designed to last up to two hours, covering all of the material for the course. It will be the same in format as the three short examinations given during the regular semester (open book, open notes), and you will be responsible for the same material on the final exam as in all three of the short exams, plus material covered after the third exam, combined. The final examination is scheduled for 5:00 to 7:00 PM Monday 10 December 2007.

(3) Calculation of grades: Each computer assignment is worth 20 points; each of the first three examinations is worth 15 points; and the final examination is worth 30 points. Grades will be calculated as a percentage of total points.

(4) Early and late assignments and examinations: I will cheerfully accept early computer assignments. I will not schedule an examination prior to the date of the exam scheduled in the syllabus. If a computer assignment is not turned in on time, for whatever reason, a new assignment will be required (same statistical procedures, different data). No extensions, no exceptions. There is no penalty to your final grade for a late computer assignment or a missed short examination. Instead, if any of the short examinations is missed, you will be required *both* to take the final examination *and* to do the fourth computer assignment. The missed examination will simply be dropped, and the final examination will be used as the makeup examination. No extensions, no exceptions other than for religious holy days (see below). Failure to take the scheduled final exam, if you are required to do so, will result in a grade of Incomplete for the course, and will need to be made up next semester.

Other Policy Issues

Now some boilerplate syllabus material. It should be unnecessary at this level, but don't make the mistake of not taking it seriously.

Attendance policy: OK, this topic is strongly recommended for inclusion in the syllabus, but you are graduate students; do I really need to tell you that in a graduate seminar, your absence will be noted and cannot possibly have a positive impact on your grade? I don't grade on attendance, but attendance *will* matter.

Tobacco products: Use of tobacco products is not allowed in the classroom at any time. The first violation of this policy will result in a warning; the second violation will result in your dismissal from the class.

Recording classes: Students are free to make audio recordings of class lectures and discussions with the understanding that no copies or transcripts of these lectures or recordings are sold or otherwise distributed for use by others. The instructor also reserves the right to record classroom lectures and discussions.

Cell phones and other electronic equipment: Please make sure that cell phones and any other electronic communication devices are set to "silent" mode during class.

Academic honesty: The faculty of the College of Criminal Justice expects students to conduct their academic work with integrity and honesty. Acts of academic dishonesty will not be tolerated and can result in failure of a course and dismissal from the university. Academic dishonesty includes but is not limited to cheating on a test, plagiarism, collusion (the unauthorized collaboration with another person in preparing work offered for credit), the abuse of resource materials, and misrepresentation of credentials or accomplishments as a member of the college. The University's policy on academic honesty and appeal procedures can be found in the manual entitled Student Guidelines, distributed by the Division of Student Services (Reference Section 5.3 of the SHSU Student Guidelines). Students are encouraged to review the statements presented in the SHSU Guidelines that refer to instances of academic dishonesty, available in the Office of Student Life or online at

http://www.shsu.edu/students/guide/studlife.html#conduct.

Disabled student policy and services for disabled students: In accordance with SHSU policy, "Students with a disability which affects their academic performance are expected to arrange for a conference with the instructor in order that appropriate strategies can be considered to ensure that participation and achievement opportunities are not impaired." The physically impaired may contact the Director of the Counseling Center as chair of the Committee for Continuing Assistance for Disabled Students by telephone (ext. 1720). Any student in the class may request special learning assistance by submitting a written statement describing your particular need. If necessary, you will be directed to the SHSU Counseling Center for additional assistance. For more detail on university policy regarding special learning needs, see the website at http://www.shsu.edu/students/guide/studlife.html#disabilities; see also http://www.shsu.edu/~vaf_www/aps/811006.html regarding disabled student policy, and http://www.shsu.edu/~counsel/sswd.html regarding services for disabled students.

Student absences on religious holy day policy: Consistent with the Texas Education Code and SHSU policy 861001, students who desire to be absent from a scheduled class in order to observe a religious holy day will present the professor with a written statement concerning the religious holy day(s). Should the religious holy day fall on a day when an assignment is due, the assignment will be due on the next class day. This is the one and only exception to the exam makeup policy.

Bottom Line

By the end of this course, I want all of you to be competent in selecting, computing, and interpreting appropriate statistics involving multiple predictors for real-world research problems. I will be absolutely delighted if I can justify giving everyone an A at the end of the course; that will mean that you did the reading, worked on the assignments, and had the good sense to ask for explanations whenever there was something you did not understand.

REQUIRED TEXTS:

Barbara G. Tabachnick and Linda S. Fidell (2007) Using Multivariate Statistics, fifth edition (Boston: Pearson/Allyn and Bacon)g, ISBN 0-205-45938-2. This is the main textbook for the course, and includes not only the statistics but also some instruction on the use of SPSS, and comparison of SPSS with other statistical packages. For most assignments, read this first.

Craig A. Mertler and Rachel A. Vannatta (2005) Advanced and Multivariate Statistical Methods: Practical Application and Interpretation, third edition (Glendale, CA: Pyrczak Publishing), ISBN1-884585-59-0. This is a good statistical workbook with more emphasis on SPSS than is found in Tabachnick and Fidell. For most assignments, read this second.

Darren George and Paul Mallery (2006) SPSS for Windows Step by Step: A Simple Guide and Reference, sixth edition (Boston: Allyn and Bacon), ISBN 0205331270. This has no real statistical content but instead is a manual for running SPSS. Think of this less as a textbook than as a dictionary to translate between English and SPSS. For each assignment, read this last, then go back and do the exercises in Mertler and Vannatta. (***Note: more recent editions are available and acceptable.)***

ALSO the following Sage monographs:

Edward G. Carmines and Richard A. Zeller (1979) Reliability and Validity Assessment (Thousand Oaks, California: Sage), ISBN 0-8039-1371-0. Good basic coverage of reliability and validity assessment, essential for anyone thinking of constructing additive scales in criminological or criminal justice research; also coverage of reliability analysis based on factor analysis.

William D. Berry and Stanley Feldman (1985) Multiple Regression in Practice (Thousand Oaks, California: Sage), ISBN 0-8039-2054-7. Well-written and with good coverage of regression diagnostics.

John Fox (1991) Regression Diagnostics (Thousand Oaks, California: Sage), ISBN 0-8039-3971-X. Some overlap with Berry and Feldman (1985) but additional material on influential cases and outliers.

Scott Menard (1995) Applied Logistic Regression Analysis (Thousand Oaks, California: Sage), ISBN 0-8039-5757-2. Includes useful information on model fit, explanatory power, and standardized logistic regression coefficients not covered elsewhere; also covers polytomous logistic regression.

Optional resources you may want to consider purchasing.

Samuel B. Green and Neil J. Salkind (2005) Using SPSS for Windows and Macintosh, fourth edition (Upper Saddle River, NJ: Prentice Hall), ISBN0-13-146597-X. For the topics this covers, it is actually a better SPSS manual than George and Mallery, but it lacks coverage of some topics, particularly logistic regression, that are included in George and Mallery. Also useful for the material on the American Psychological Association writing style.

James Alan Fox, Jack Levin, and Michael Shively (2002) Elementary Statistics in Criminal Justice Research, second edition (Boston: Allyn and Bacon), ISBN 020533373-7. This is a reasonably good elementary statistics textbook; you should already know this material, but this can help fill in any gaps. Chapter 9 is particularly useful as a review of nonparametric statistics, an often neglected topic in elementary statistics courses.

TWELVE FUNDAMENTAL CONCEPTS IN STATISTICS

1. CASE. One of the objects (person, place, or thing) we are observing or measuring. When human, cases are sometimes called *subjects;* subjects in survey research are sometimes called *respondents*.

2. VALUE. A symbol, usually a number, representing a measured or observed or predicted characteristic of a case.

3. VARIABLE. A characteristic of the case (or cases) we are observing or measuring, which has two or more values.

4. CONSTANT. A characteristic of the case (or cases) we are observing or measuring which has only one value, the same value for all cases.

5. LEVEL OF MEASUREMENT. The level of measurement of a variable is a way of telling how much information the values of the variable convey.

6. POPULATION. All of the cases in which we are interested.

7. SAMPLE. Some of the cases in which we are interested.

8. DISTRIBUTION. A set of values for one or more variables, for a real or hypothetical population or a sample, typically represented in a graph or a table.

9. PROBABILITY. How probable or likely it is that something (usually a particular value or set of values of one or more variables) will occur or be observed; sometimes called a *likelihood.*

10. ERROR IN PREDICTION. The discrepancy or difference between what we expect or predict and what we observe; often measured as a function of the difference between predicted and observed values of a variable (most often, that difference squared).

11. CENTRAL TENDENCY. The single predicted value of a variable which, if predicted for every case, minimizes the total (sum or product) of the errors in prediction. The central tendency is also called the *average* when referring to a single variable.

12. DISPERSION. A measure of how narrowly or widely the observed or predicted values of a variable are spread around a particular value of that variable (usually an observed or predicted or expected measure of central tendency of that variable.)

All of statistics is based on these twelve concepts.

CJ 742: Advanced Statistics I: COURSE OUTLINE

Date	Reading Assignment	Topics
		Unit I: Overview and review
Aug 22 (Week 1)	(none)	Diagnostic examination. Basic concepts and definitions; causality.
Aug 29 (2)	Tabachnick & Fidell 1-2, App. B: **skim** Tabachnick and Fidell 3 Mertler and Vannatta 1-2: **skim** George and Mallery 1-5 (Green and Salkind 1-23, 24-25, 40: **skim**)	Lab session. Overview of multivariate statistical techniques; review of levels of measurement; univariate and bivariate statistics; hypothesis testing and measures of association; contingency table analysis; bivariate linear regression and correlation.; analysis of variance (ANOVA); brief review of SPSS.
Sep 5 (3)	Tabachnick and Fidell 3 (again); focus on pp. 37-55 (sections 3.2-3.4) Mertler and Vannatta 4 George and Mallery 12-13 and 17 (Green and Salkind 38-44) Also recommended: Fox et al. (2002) *Elementary Statistics* 9 (for nonparametric statistics)	One-way and factorial within-subjects (independent samples) analysis of variance and related nonparametric techniques; one-way repeated (matched samples) analysis of variance and related nonparametric techniques; post-hoc tests and planned contrasts in ANOVA. Bonferroni adjustments. ***Note: in one way or another, we will be returning to the material on ANOVA in every part of this course, so be sure you know this really well the first time.***
Sep 12 (4)	Tabachnick & Fidell 4 Mertler and Vannatta 3	Lab session. Data screening and transformations.
		Unit II: Analysis of interval/ratio dependent variables: Multiple regression and ANOVA/ANCOVA
Sep 19 (5)	Tabachnick & Fidell 5 Mertler and Vannatta 7-8 George and Mallery 16 (Green and Salkind 33)	Ordinary least squares multiple regression analysis; introduction to path analysis. **Assignment 1: Review of univariate and bivariate statistics**
Sep 26 (6)	Berry and Feldman (1985) all Fox (1991) all, including appendix George and Mallery 28 except p 360-363 **Review** Tabachnick and Fidell 4 **Review** Mertler and Vannatta 3	Lab session. Regression diagnostics; testing regression assumptions; model building and model selection in multiple regression analysis; introduction to advanced topics in regression analysis.
Oct 3 (7)	Tabachnick & Fidell 6 Mertler and Vannatta 4-5 George and Mallery 12-14 (Green and Salkind 26)	Analysis of variance and covariance (ANOVA and ANCOVA); modeling ANOVA and ANCOVA using regression analysis and dedicated ANOVA routines. **Assignment 2: Analysis of interval/ratio/continuous dependent variables**
		Unit III: Analysis of categorical outcome variables: Logistic regression and related techniques
Oct 10 (8)	Menard (2002) 1-3, appendix Tabachnick et al. 10 Mertler and Vannatta 11 George and Mallery 25	Lab session. Logistic regression analysis for dichotomous dependent variables; parallels with multiple regression and ANOVA.

Date	Reading Assignment	Topics
Oct 17 (9)	Menard (2002) 4-5 Tabachnick and Fidell 16: **skim** George and Mallery 26-27: **skim** George and Mallery pp. 360-364: **skim**	Logistic regression diagnostics; testing logistic regression assumptions; model building and model selection in logistic regression analysis; logistic regression analysis for polytomous dependent variables; logistic regression and loglinear analysis.
Oct 24 (10)	Tabachnick & Fidell 12: **skim; *read first*** Tabachnick and Fidell 9: ***read second*** Mertler and Vannatta 10 George and Mallery 22 (Green and Salkind 34)	Lab session. Canonical correlation and discriminant analysis. *(Note: canonical correlation analysis actually begins the section on models with multiple outcomes.)* **Assignment 3: Analysis of nominal/ordinal/categorical dependent variables**
		Unit IV: Measurement
Oct 31 (11)	Carmines and Zeller (1979) 1-4 George and Mallery 18 (Green and Salkind 36-37) Tabachnick and Fidell Appendix A (**Read this last**)	Lab session. Reliability and validity assessment; repeated measures ANOVA and related techniques; Cronbach's alpha and additive scaling; handling missing data in additive scales; introduction to matrix notation.
Nov 7 (12)	Tabachnick & Fidell13 Mertler and Vannatta 9 Carmines and Zeller (1979) Appendix George and Mallery 20 (Green and Salkind 35)	Principal components and exploratory factor analysis and scale dimensionality; reliability assessment using principal components and exploratory factor analysis. **Assignment 4: Additive scaling and factor analysis**
		Unit V: The general (and generalized) linear model
Nov 14 (13)	NO CLASS (but still assigned reading) Tabachnick & Fidell 7 Mertler and Vannatta 6 George and Mallery 23 (Green and Salkind 27)	AMERICAN SOCIETY OF CRIMINOLOGY MEETING Multivariate analysis of variance and covariance (MANOVA and MANCOVA).
Nov 21	NO CLASS	THANKSGIVING BREAK
Nov 28 (14)	Tabachnick and Fidell 8 George and Mallery 24 (Green and Salkind 28-39)	Other (besides reliability analysis) applications of repeated measures MANOVA and MANCOVA **Assignment 5 (optional): Analysis of multiple dependent variables (*real* multivariate analysis)**
Dec 5 (15)	Tabachnick and Fidell 17	Lab session. The general linear model: overview and summary
Dec 12	(Official final examination date)	Final examination/Final assignment due