PSYCHOLOGY 204B:

CAUSAL MODELING OF CORRELATIONAL DATA

Room: 192 Young. *Times*: TR 2:10-4:00.
Webpage: http://psychology.ucdavis.edu/Simonton/p204bwmain.html

Instructor: Dean Keith Simonton
Office: Young 102D. *Office Hours*: TR 4:10-5:00 (or by appointment).
Phone: 752-1677. *E-mail*: dksimonton@ucdavis.edu

TA: Joel S. Steele
Office: 266 Young. *Office Hours*: TBA (or by appointment).
Phone: 754-5370 *E-mail*: jssteele@ucdavis.edu

1. COURSE GOALS: To provide you with professional competence in applying causal modeling methods to correlational data in the behavioral sciences. The emphasis is testing rival models using correlations among observed variables. Because computers almost invariably execute these techniques, you will receive elementary instruction in how to implement the analyses on standard statistical software. The emphasis is on application rather than mathematical foundations, and therefore algebraic derivations are used solely to encourage you to obtain a conceptual understanding of the techniques.

2. ENTRY LEVEL: Course 204a, or consent of instructor. Designed primarily for graduate students in the behavioral sciences who are in the first or second year of training. Upon the instructor's approval, undergraduates who have taken an upper-division class in statistics may take the course as well. Please note that university-level mathematical and computer competence is assumed. If you do not feel comfortable with the use of algebra or the installation and execution of computer software, you may face difficulties that will interfere with your ability to do well in the course.

3. TOPICAL OUTLINE: The course is divided into three parts:

A. Multiple Regression/Correlation Analysis – Bivariate correlation and regression methods are generalized to accommodate multiple independent variables. Partial and semipartial correlations, partial regression coefficients, and multiple correlation are all introduced as responses to the "third variable" problem in causal inference. Also discussed are the issues of inflated R^2, multicollinearity and suppression effects, and standardized versus unstandardized parameters. This section concludes with an introduction to executing multiple regression/correlation analyses using commercially available computer software.

B. Advanced Techniques and Topics – Other issues in multiple regression will be discussed, as time permits. These include (a) nominally scaled dependent and independent variables, (b) interaction effects and curvilinear functions, (c) the consequences of outliers, autoregressive residuals, and violations of multivariate normality, and (d) missing data. Incorporated throughout this section is a discussion of how the results are normally presented in the professional literature.

C. Path Analysis and Structural Equation Models – An introduction to the basic principles of formulating and testing causal models using systems of linear equations. Discussion of how to use covariance algebra, the basic theorem, or the tracing rule to work out the implications of alternative linear models. Although the focus will be on equation-by-equation ordinary least squares parameter estimation, an overview will be given of such system estimation procedures as generalized least squares and maximum likelihood. In addition, as time permits, we will discuss (a) model misspecifications, especially measurement errors, correlated disturbances, and reciprocal causality or feedback loops, (b) the identification problem, criteria for identification, and the testing of underidentified models, (c) nominally scaled endogenous and exogenous variables, (d) multiple indicators in latent variable models, and (e) the use of such specialized software as EQS and LISREL.

4. GRADING AND COURSE REQUIREMENTS: There are three equally weighted take-home exams that follow each section of the course. Also, scattered throughout the quarter, will be some assignments that will be graded. In addition, on the last day of class, students are required to submit a research project; this requires that a causal analysis be applied to appropriate data of the student's own choosing; the output of a computer analysis must be interpreted and written up in standard journal format. If the student has no data set on which to conduct an appropriate analysis, the instructor will provide a data set (reluctantly).

The three exams are worth 25% each, the research project 15%, and the assignments 10%. All exams and assignments must be turned on time for full credit. Otherwise, one-third of a grade will be deducted each day one is late.

Please note that I assume that *all* exams and homework represent *individual* work. Therefore, collaboration on the take-home exams or homework assignments will be considered violations of the student code of ethics (see http://sja.ucdavis.edu/pdf/collab.pdf).

5. READING: The following textbook will be used:

Jacob Cohen, Patricia Cohen, Stephen G. West, & Leona S. Aiken.
Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.).
Mahwah, NJ: Lawrence Erlbaum, 2003.

This is a much-expanded version of a classic text. First published in 1975, it went through a second edition in 1983. Given that track record, it will provide a standard text for years to come.

6. SCHEDULE:

Part A: *Multiple regression/correlation analysis* (begin on 1-6).
Readings from: Cohen et al.: Chapters 1-3, 10 (selections).
Exam 1: Get on 1-29 and return on 1-30 (by 4 pm).

Part B: *Advanced techniques and topics* (begin on 2-5).
Readings from: Cohen et al.: Chapters 4-9, 10-11 (selections).
Exam 2: Get on 2-19 and return on 2-20 (by 4 pm).

Part C: *Path analysis and structural equation models* (begin on 2-26).
Readings from: Cohen et al.: Chapter 12 (selections).
Exam 3: Get on 3-13 and return on 3-14 (by 4 pm).

Project: Due on 3-13 (last day of class).