#131, Spring 2007

GRADUATE SCHOOL OF SOCIAL WORK AND SOCIAL RESEARCH
BRYN MAWR COLLEGE

DATA ANALYSIS
#131
Spring Semester 2007

Instructors
Thomas P. Vartanian (tvartani@brynmawr.edu)
Linda Houser (lhouser@brynmawr.edu)
Suzanne McMurphy (smcmurphy@brynmawr.edu)

Teaching Assistants
Sandra Guillory (sguillory@brynmawr.edu)
John McWilliams (jmcwilliam@brynmawr.edu)
Lisa Reedich (lreedich@brynmawr.edu)

Web Page Address: http://www.brynmawr.edu/Acads/GSSW/Vartanian/
Then click on: Data Analysis, SW131

Course Description

This course prepares students to seek answers to fundamental questions about social work practice, social policy and social programs through the analysis of quantitative data. Students gain an appreciation for the use of statistics in evidence-based practice through an emphasis on the application of key statistical techniques and interpretation of results so that students can utilize statistics appropriately in their practice. Particular attention is paid to the legitimate and illegitimate use of statistical techniques and the misuse of data in support of discriminatory positions and practices toward vulnerable and at-risk populations. Students become critical consumers of research articles through familiarization with statistical methods used to examine issues related to poverty and other social and economic welfare issues.

Course Objectives

Students will be expected to demonstrate the following competencies:

1. ability to use the computer, statistical software, and basic statistics to perform univariate, bivariate, and simple multivariate analyses of quantitative data;

2. ability to design and execute an original analysis of quantitative data;

3. ability to select appropriate descriptive statistics;

4. ability to generate and test hypotheses;

5. enhanced skill in writing a research article;

6. awareness of the ways in which statistics can be misapplied or misinterpreted to permit false conclusions, often in support of erroneous or biased positions with respect to women and racial, ethnic, vulnerable and at-risk groups;

7. ability to comprehend, effectively criticize, and utilize research articles to inform their professional practice;

8. ability to engage in dialogues with social science researchers about interpretations of current research findings and important directions for further inquiry;

9. ability to conduct and interpret research on issues related to poverty, social welfare, and social
justice;

(10) interpret statistics as they relate to students’ individual practices

(11) Understand that different populations may respond differently to particular conditions and that these variations need to be incorporated into research models

(12) Understand that generalizations cannot be made across people that differ in their sexual orientation, gender, race, ethnicity, and social class.

Class Policies

Students are expected to attend all class sessions and submit written assignments on time. The instructor should be notified in advance if the student expects to miss a class meeting or assignment due date. Class attendance is a routine expectation, and it is assumed that students will take an active role in class discussions.

Late submission of assignments must be negotiated in advance with the instructor. Failure to negotiate a late submission of an assignment will mean the loss of points for the final grade. Also, if assignments are handed in past the time of the negotiated data, points will be deducted for the final grade. It is assumed that all written and computer work will be done independently, unless otherwise specified by the instructor. Students may help each other with their computer assignments and in their analysis of data for their final research project. However, students may not do the same research project together or work on similar research projects together. Students who do not undertake independent work on the research paper will fail the class. All written work must be produced in MS Word.

Grades for this course are “Satisfactory” and “Unsatisfactory”, and include written and computer assignments, as well as class participation, in accordance with School policy. In order to achieve the intended outcomes for the course, the student must complete all work satisfactorily. Students who miss more than two class sessions will not pass the class unless they have discussed the reasons and received approval for such absences with their instructor.

Grades for the class will be determined on the basis of three factors (with the proportion of your final grade for each factor):

Mid-term exam: 40%, Final Exam: 40%, Paper: 20%

Students must hand in all computer assignments and are expected to participate in class discussions.

Students may not proceed to the next semester or summer semester until all work from the previous semester or summer semester is complete or until a grade of Permanent Incomplete or Unsatisfactory is issued.

Assignments

Students are expected to complete assigned readings in advance of class meetings. In addition to the required readings listed in the course outline below, other reading assignments will be made from time to time.

Computer exercises will be provided to familiarize students with basic statistical tools in both SPSS and Microsoft Excel.

There will be a formal midterm and final exam.

Each student will develop and conduct an original analysis of data from the 1992/1999/2003 Panel Study of Income Dynamics (PSID), the Child/Adolescent sample, or the 1997/2002 Child Development Supplement. The data will
be available on the computers in the computing lab in the school. All three of these data sets originate from the Survey Research Center at the University of Michigan. We have extensively manipulated the data sets so that they contain variables of interest to social work students. The data sets were collected and are displayed in such a way so as to not identify those who partook in the sample. Value judgments were made by those at the Survey Research Center at the University of Michigan, who collected the data for the PSID, as to which groups would be oversampled so as to provide researchers with adequate sample sizes to better understand factors affecting particular groups. For this reason, African American and poor households were oversampled in the data so as to get relatively large sample sizes for these groups. The data is then weighted to make it representative of the U.S. population.

The first two data sets explained here come from the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal data set that started in 1968 and has continued for 32 years, with the last version available for 2003. The PSID is a representative data set of the U.S. population when weighted (the sample is a disproportionate stratified sample -- we'll talk more about weighting later). The first data set, the Panel Study of Income Dynamics 1992/1999/2003, includes over 4,000 individuals from the United States. I have included information for a group of individuals that were in the sample in 1992, 1999, and 2003. The data include information on such characteristics as income, race, education, economic situation while growing up, health status, religious preferences, housing status, whether the person was born with low birth weight, and government income assistance. The variables from 1999 contain information on such health conditions as whether the individual has cancer, heart disease, has ever had a stroke or a heart attack, has mental loss or emotional problems, or has a learning disability, and the number of years the person has had these conditions. You will be able to determine the relationship between variables for the entire sample or for particular subgroups (such as the elderly, the poor, single mothers, people of different races, husbands, wives, single adults, or those with little education). The PSID is used by a great number of researchers throughout the country because it provides a wealth of information on a nationally representative group of individuals and families.

The Child/Adolescent Sample comes from the 1968 to 1997 panels of the PSID. We examine variables for children aged 10-14 (generally – some variables are from when the child was born) and then examine these same children when they become adults. For example, we examine such childhood/adolescent variables as parental level of education, whether parents are married or not, marital status of the child’s mother when the child was born, birth weight, income, wealth, and government assistance. Also included in the data are adult variables such as age when first married, number of marriages, number of children, health status, religious preference, level of education, income, occupation, wealth and government assistance received (such as welfare). From this data, you can determine which childhood/adolescent factors help predict adult outcomes. There are nearly 4,000 individuals in this child/adolescent sample.

The third data set is called the Child Development Supplement (CDS). It is a supplement to the Panel Study of Income Dynamics. These data contain information on children, aged 0 to 12 in 1997 and ages 5 to 17 in 2002, and their primary caregivers. From these data, you can help determine what factors affect children’s health, behavior, cognitive ability, and a great variety of other outcomes. For example, you can determine how family circumstances while growing up, such as the mother’s emotional state, parental expectations of the child, parental disciplinary practices, whether the child comes from an immigrant family, the level of cognitive stimulation for the child, child’s perception of self, family income, perceived neighborhood safety, race, wealth, family composition, welfare receipt, religious practices, and many other variables, affect how children fare. Some of the principle outcome variables for these data include reading and math abilities, both internal and external behavioral problems as well as overall behavioral problems, health issues, including a great variety of problems, and many other possibilities.

Warning: All data used in this class are the property of Professor Thomas Vartanian. Under no conditions can any student claim title to any of these data or publish without the written consent of Professor Vartanian.

Each student is advised to take a close look at the codebooks for each of the data sets to get a more complete picture of what each set of data contains. Students will use one of these three data sets for the rest of the semester. Each
student will select variables to use in hypothesis testing and use SPSS for Windows to execute appropriate analyses of the relationships between these variables. Three written assignments will be produced with particular objectives:

#1  A two to three page paper containing a problem statement, significance for social work, and the hypothesis to be tested. This assignment is due in the fourth week of class. The objectives for this first paper include:
   a. Listing of the independent, dependent, and control variables and their SPSS names.
   b. Listing of the level of measurement of each of these variables.
   c. Listing of the hypotheses to be tested
   d. Explaining the significance of the research to social work, in general, or to specific fields of social work

#2  A two to four page paper including a statement of the research problem and a review of the relevant literature. This assignment is due in the eighth class session. The objectives of this second assignment include:
   a. Understanding what factors others have found to affect the outcome variable.
   b. Determining whether there are differences in outcomes for particular groups, especially vulnerable groups
   c. Determining which factors, if any, have not been used to examine the outcome variable.

#3  A term paper including a statement of the research problem, a review of relevant literature, a description of the data and research methods, presentation of the findings, and a discussion of the implication of the findings for social work practice. This paper should be written in the form of a journal article (10-12 pages plus a bibliography). This assignment is due in the 12th class session. The objectives of this third assignment include:
   a. Understanding and accurately describing the results from the statistical analyses
   b. Linking the results to the findings in published research
   c. Understanding how variables left out of the analysis may affect the final results

All three of these assignments are explained in further detail on the web page for the class.  
http://www.brynmawr.edu/Acads/GSSW/Vartanian/PaperAsg.htm

Texts


Topic Outline

<table>
<thead>
<tr>
<th>Session Number</th>
<th>Dates</th>
<th>Topic and Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan. 22-24</td>
<td>INTRODUCTION AND OVERVIEW OF THE COURSE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DESCRIBING AND ANALYZING DATA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FREQUENCY DISTRIBUTIONS AND GRAPHS</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Readings:</strong> Weinbach and Grinnell, Chapters 1 and 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Computer assignment #1, available on the web page for the course.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Due in week 2.</td>
</tr>
</tbody>
</table>
2 Jan. 29-31 CENTRAL TENDENCY AND DISPERSION
Measures of central tendency include the mean, median and mode. Measures of dispersion include the standard deviation and the variance.

Readings: Weinbach and Grinnell, Chapter 3

Computer assignment #2, available on the web page for the course. Due in week 3.

3 Feb 5-7 NORMAL DISTRIBUTIONS AND Z-SCORES
We will be analyzing how hidden family factors may affect poverty among African American teenagers.

Readings: Weinbach and Grinnell, Chapter 4 and 6

4 Feb. 12-14 INTRODUCTION TO HYPOTHESIS TESTING
The logic of hypothesis testing
The meaning of statistical significance

Readings: Weinbach and Grinnell, Chapters 5 and 7

Due: Written assignment #1 (2-3 page paper) due

5 Feb. 19-21 CROSS-TABULATION AND CHI-SQUARE
We will learn how to conduct significance tests when using two nominal scale variables. We will use chi-square analysis to compare groups of sexually abused and non-abused children.

Readings: Weinbach and Grinnell, Chapter 10

Computer assignment #3, available on the web page for the course. Due in week 6.

6 Feb. 26-28 WRITING AND EVALUATING RESEARCH REPORTS
Important components
Criteria for evaluating reports
Narrative presentation of data
Table construction
Presenting information in graphs and charts
Appropriate representation of subgroups as well as accurate representation of subgroups.

Readings: Weinbach and Grinnell, Chapters 17 and 18.

7 March 5-7 MIDTERM EXAM
March 12-14 SPRING BREAK

8 March 19-21 CORRELATION AND SIMPLE REGRESSION
We will learn how simple regression and correlation are used to study issues related to poverty and social welfare. Examples from the Panel Study of Income Dynamics and the Child Development Supplement to the PSID will be used to examine factors affecting poverty for ethnic minority populations and women.

Readings: Weinbach and Grinnell, Chapters 8 and 9

Computer assignment #4, available on the web page for the course. Due in week 9.

Due: Written assignment #2 (2-4 pages. Research questions, hypotheses, and literature review.)

9 Mar. 26-28 MULTIVARIATE OPTIONS
We will examine how bivariate regression and correlation may misrepresent relationships. We will learn how to use control variables to partial out factors and find true relationships. We will also learn how to use nominal scale variables as independent variables within regression models. From the PSID and CDS, we will examine the differential effects of economic conditions (such as state level unemployment rates and welfare availability) and family factors (such as marital status) on minority populations and women.


Computer assignment #5, available on the web page for the course. Due in week 10.

10 April 2-4 LOGISTIC REGRESSION
We will focus on how to use regression analysis with a nominal scale dependent variable. We will examine such issues as how childhood factors affect poverty status as an adult for different groups, including African Americans, Whites, men and women. We will also focus on how condom availability affects the likelihood of sexual activity for kids in
high school in New York and Chicago. We will learn about how we can and cannot generalize these results.


Computer assignment #6, available on the web page for the course. Due in week 11.

11 April 9-11 LOGISTIC REGRESSION
Continuation of logistic regression focusing on interpretation of results to various at-risk and vulnerable populations


12 April 16-18 GROUP T-TESTS AND ANALYSIS OF VARIANCE

We will learn how to conduct research on groups as independent variables, for interval/ratio scale dependent variables. We will be examining issues related to differences in racial, ethnic and gender groups, for factors such as income, education and welfare. We will also examine how people of different racial groups assess whether a child has been maltreated or not.

Due: **Written Assignment #3, Term paper**

**Readings:** Weinbach and Grinnell, Chapter 11

Computer assignment #7, available on the web page for the course. Due in week 13.

13 April 23-25 GROUP T-TESTS AND ANALYSIS OF VARIANCE and REVIEW
#131, Spring 2004

14 April 30-May 2 FINAL EXAM