Quantitative Methods in Applied Anthropology

Course: ANG 5486 (Fall 2008); THR, 9 am – 12 pm & 2 pm – 5 pm

Instructor: Christian Wells, cwells@cas.usf.edu, 813/974.2337

Office hours: WED, 10 am – 12 pm; FRI, by appointment

Most social science research requires the recognition and analysis of *patterns* in empirical data. However, these are rarely immediately apparent in the large batches of numbers obtained during research. Quantitative methods of data analysis provide an indispensable means of detecting such *patterning* in a systematic and organized fashion.

÷

This course is intended as an introduction to quantitative methods and reasoning for the anthropologist, although other social scientists make use of the same principles and techniques we will examine. Most of the techniques are fairly standard, deriving from either "classical statistics" developed in the early 1900s or from the more recent "exploratory data analysis" approach. The course focuses on topics covered in *Statistics for Anthropology* (L. Madrigal, 1998, 2nd printing, Cambridge University Press) and *Quantifying Archaeology* (S. Shennan, 1997, 2nd edition, Edinburgh University Press), occasionally supplemented by additional readings that present case studies of the methods we will discuss in class. In addition, we will cover the basics of using computers for the analysis of anthropological data with packaged statistical software, primarily SPSS.

÷

The main objectives of the course are:

- to introduce you to the quantitative analysis of data using computers
- to help you to become a thoughtful and informed consumer of the quantitative anthropological literature
- to discuss in detail the sorts of problems that can be addressed quantitatively in anthropology
- to provide you with the background necessary for designing research to yield data that may be easily and productively analyzed using quantitative methods

Format, Requirements, & Evaluation

The class meets once a week, but is divided into two parts. The first part, which lasts about an hour, consists of an illustrated lecture and group discussion of the weekly topic. You are uniquely responsible for making certain that you are prepared for class, which includes having completed a thorough examination of the assigned readings and, when available, looking over the relevant examples. Take notes. Come prepared to ask questions.

The second part is devoted to a workshop-style examination of an anthropological problem using the methods we have learned in class that day. This is an in-class work session; attendance is absolutely mandatory without exception. Please be aware that you need to bring a portable laptop computer (preferably with a wireless Internet connection) with the latest versions of Excel and SPSS installed. Please see the instructor if you do not have access to a computer or the appropriate software. Since adequate coverage of each problem requires more time than we have scheduled for class, you will be responsible for completing the assignment. Each assignment will be due at the start of the next class; late papers will not be accepted under any circumstances other than a fully documented medical or criminal emergency.

There will be nine graded assignments (each worth 10 percent of your final grade), which require the application of a variety of quantitative techniques to anthropological data (note: there are three additional assignments, but these will be completed in class and will not count toward your final course grade). Evaluation of your performance on each is based on the quality of your analysis of the data and the quality of the argument you make concerning the anthropological question(s). You should be thorough in answering all questions posed for each problem, and present your results in a brief essay that references supporting figures and tables, as well as any relevant publications. Essays need to be typed, and the write-up is not to exceed five double-spaced pages of text using 12-point TNR font with 1-inch margins. Figures, tables, and any calculations may be appended as necessary. Such illustrations need not be of publication quality, but they need to be clear and well-labeled.

For each essay, you should briefly state the problem and the methods used to solve it; you may not assume the reader's knowledge of the methods we have discussed. Other than that, you should present the results as you would in a peer-reviewed article. In particular, you need to be very clear about what procedures have been employed (e.g., don't say "correlation" if you mean the "product moment correlation coefficient" or don't say "factor analysis" if you mean "principal components analysis"), the set of data analyzed along with any transformations (trimming, z-score conversions, etc.), exactly what is shown quantitatively, and exactly what of substance you conclude. You may wish to coordinate your work times with other members of the class so you can discuss any problems that arise. While I encourage you to discuss the methods and issues and help each other with computer problems, you will need to execute the procedures and write up the essays independently.

Toward the end of the semester you will use what you have learned about quantitative methods to summarize and critique a professional article from the recent (post-1990) anthropological literature. Your summary and critique will take the form of a 15-mintue class presentation, worth 10 percent of your final grade. Article selection is up to you, but must be approved by me at least two weeks in advance of your presentation.

Grading Policy

Performance evaluation includes the following grades: A (“excellent”), B (“good”), C (“average”), D (“poor”), F (“failure”), and I (“incomplete”). These grades are earned based on the following scale: A+ = >97, A = 92-97, A- = 90-91, B+ = 88-89, B = 82-87, B- = 80-81, C+ = 78-79, C = 72-77, C- = 70-71, D = 60-69, F = <60. Please note that incompletes (“I” grades) will not be issued after November 1, 2006 and only then at the discretion of the instructor. The USF College of Arts and Sciences dictates that incomplete grades should only be granted when, due to circumstances beyond the control of the student, only a small portion of the required work remains undone and the student is otherwise passing the course. Students seeking an “I” grade will be required to sign a contract with the instructor, specifying the work to be completed and the deadline for completion.

Schedule

Date	Topic	Madrigal	Shennan	Wells
31-Aug	Approaches to Quantitative Methods in Anthropology: Statistics and EDA; Exercise #1 (completed in class)	1	1 & 2	
7-Sep	Pictorial and Numerical Summaries of a Single Variable; Exercise #2	2 & 3	3 & 4	2004a
14-Sep	Pictorial and Numerical Summaries of Multiple Variables; Exercise #3	2 & 3	3 & 4	2002
21-Sep	Probability, Statistical Inference, and Random Sampling; Exercise #4	4 & 5	5, 6, & 14	2005
28-Sep	Basic Parametric Tests: Student's *t* and Chi-squared; Exercise #5	6 & 11	6 & 7	2003
5-Oct	Analysis of Variance: Univariate and Multivariate Methods; Exercise #6	7	5	2007a
12-Oct	Basic Non-parametric Tests: Mann-Whitney *U* and Kruskal-Wallis *H*; Exercise #7	8	5	
19-Oct	Linear Correlation: Pearson's *r* and Regression Analysis; Exercise #8	9 & 10	8, 9, & 10	2006
26-Oct	Cluster Analyses: Hierarchical and Partitioning Methods; Exercise #9		11	2007b
2-Nov	Multivariate Analyses: Principal Components Analysis and Discriminant Function Analysis; Exercise #10		12 & 13	2000
9-Nov	Spatial Analyses: Geostatistical Interpolation and Kriging; Exercise #11 (completed in class)			2007c
16-Nov	Simulation Analyses: Predictive Modeling and Monte Carlo Simulation; Exercise #12 (completed in class)			2004b
30-Nov	Student Presentations			
7-Dec	Student Presentations			

THINGS TO REMEMBER ABOUT WRITING ABOUT QUANTITATIVE ANALYSIS

• Look at your data first using simple tables and pictures. Often this tells you everything important. If not, it will tell you what is sensible or not sensible to do next.

• Statistical analysis is not a way to arrive at certainty; it is a powerful aid in discerning what your data suggest and how strongly they suggest it. This is often done better by an estimation approach than by hypothesis testing.

• If you must do a hypothesis test, report the actual probability level obtained and don't treat some arbitrary level, such as 5%, as a talisman that tells you what to think.

• It's not the sampling fraction that matters; it's the size of the sample.

• Proportions, percents, and ratios represent something relative to something else. They are fractions, with a numerator and a denominator. When you write, always report the denominator. When you read, always ask yourself whether you understand what denominator is implied.

• "Frequency" should always mean count of something, rather than ratio of something to something else; the latter should be referred to as “relative frequency.”

• “Data” is plural while “datum” is singular; thus, we say, “these data are” instead of “the data is.”

• “Dataset” is one word (not “data set” or “data sets”).

• The character, %, is usually spelled out, percent; thus, “10%” should be “10 percent.”

• Use a space to separate the unit of measure from the value of the case; thus, “10ml” should be “10 ml.”

• Spell out whole integers less than 10; use Arabic numerals for whole integers greater than nine (unless they appear first in a sentence, then spell them out): zero, one, two, three, four, five, six, seven, eight, nine, 10, 11, 12, 13, 14, 15, etc., but "Fifty-seven percent of the people are tall." An important exception: do not spell out values that have measurements; thus, “five square meters” should be “5 m^2.”

• “Boxplot” is one word (not “box plot”) and is a noun, so there’s no need to qualify it, i.e., “boxplot chart;” an alternative reference is “box-and-whisker plot.” Similarly, “ANOVA Analysis” or “PCA Analysis” is unnecessarily repetitive.

• If most of the data are to the left in a histogram and there are outliers to the right, we say that the data are “skewed right,” that is, we note where the skew occurs.

• When comparing histograms, they must be based on the same scale for both X and Y axes.

• Lowercase “n” reports the sample size, while uppercase “N” reports the population size; also, there is no need to italicize “n” or “N” because they are not statistics (i.e., they are not calculations).

• All tables and figures should be numbered and have descriptive captions; the labels (e.g., “Figure 1” or “Table 2”) should be referenced in the text when you discuss each table or figure.