Spring 1-2016

ANTY 408.01: Advanced Anthropological Statistics

Randall R. Skelton
University of Montana - Missoula, randall.skelton@umontana.edu

Recommended Citation

Skelton, Randall R., "ANTY 408.01: Advanced Anthropological Statistics" (2016). *Syllabi*. 4627.
https://scholarworks.umt.edu/syllabi/4627

Advanced Anthropological Statistics
ANTHROPOLOGY 408: Spring 2016

TR 8:10-9:30   SS 258

Dr. Randy Skelton   **226 Social Sciences Building**
Office Hours: **MWF 8:00-8:50, TR 10:00-11:00**
Phone: 243-4245   **Email: randall.skelton@umontana.edu**

GOALS AND OBJECTIVES

The goal of this class is to learn several advanced (multivariate) methods of data analysis and to learn the skill of writing a scientific paper. The focus will be on use of statistical software to perform analyses, with interpretation and write-up of the results obtained. Students who pass this class will:

- Learn to use several types of statistical analysis including multiple regression, principal components analysis, cluster analysis, discriminant analysis, and more.
- Explore how these analysis can be applied to novel situations by carrying out a project that involves the use of data analysis.
- Gain facility with a statistical software packages such as SPSS.
- Build the ability to interpret the results of multivariate statistical analyses and express them in a professional manner.
- Become familiar with standard scientific paper style and format.
- Gain experience with finding sources through library databases.
- Come to appreciate the vast array of data that is available on the web.

ADMINITRIVIA

Moodle Supplement

There will be a Moodle supplement for this class, where I will post various types of useful materials and information, including required materials. The people at IT Central in SS120 and Moodle Tech Support (243-4999, umonline-help@umontana.edu) can help you with access and technical issues. As your instructor I can only be responsible for <u>content</u> placed on Moodle – not for its administration or technical issues.

Required Materials

Text: Landau, Sabine and Everitt, Brian S., 2004. *A Handbook of Statistical Analyses using SPSS.* Chapman & Hall/CRC Press. Hereafter I will refer to this text as "the Handbook". The Handbook provides a walk-through of many of the methods we will be covering. The Handbook will be most useful to you when you are doing your assignments, and need not be read before coming to class.

Online Resources: For each week I have some required browsing listed. Some of this is for help with your assignment. There are many statistical texts online, some of which I have links to in the "Helpful Materials" section of the class Moodle supplement. The most useful of these materials are made available on the WWW by Statsoft Inc, Karl L. Wuensch (a professor in psychology at East Carolina University), and William K. Trochim (a professor in policy analysis and management at Cornell University).

Statistical Package: We will use SPSS. SPSS is available in the Fred W. Reed Social Sciences Research Lab (SSRL) and other campus computing labs. You may also buy SPSS.

Other Software: I assume that you have, and know how to use, Microsoft Office products, especially Word and Excel. You may also download the free office package OpenOffice and use it instead, though I don't guarantee that it operates exactly the same as MS Office.

Computer access: You will need access to a computer with SPSS installed. SPSS is installed on computers for student use in the SSRL. We will have an orientation to the use of these labs by the SSRL staff early in the semester. Also, the computers in the UC 225 general student lab are supposed to have SPSS installed. You will need to show your GrizCard for access to the general student labs.

Data Storage: You will need some mechanism for storing the data sets you use and the output from the statistical software. The best option for this is a USB flash drive (also known as a memory stick, pen drive, flash drive, etc.), though UMBox is probably a reasonable alternative.

How will this class work?
1. First 2/3 of the semester. We will explore several methods of advanced statistical analysis. The focus will be on using SPSS to perform the analysis, interpreting the output, and writing up the procedure in standard scientific paper format. We will meet at every normally scheduled class meeting time. Each week there is an assignment due and you will be expected to do the analysis requested, write up your results, and submit them to me by uploading them through Moodle. Moodle doesn't have a way for me to send materials back to you, so I will do that using email to your official University email address (what I see in Moodle). You will need to either check your University email regularly or forward it to where you normally check your email. Most weeks there will be a lecture on Tuesday, and we will work with data on Thursday.
2. Last 1/3 of the semester. You will each do a project in which you analyze a data set of interest to you in order to draw some conclusions about some topic of interest to anthropology. Grad students should use the data set they are working with in their thesis or dissertation research, if possible. We will continue to meet for class, and I will use this time to explore and demonstrate additional statistical and analytical methods. I will not allow you to fall behind or put off the steps of the project until the end, and there is an assignment related to your project due every week.

Grading
For undergraduate students, your grade will be based on attendance, preparation, and participation (25%); weekly exercises you complete (30%); and your project (45%).   For graduate students, your grade will be based on attendance, preparation, and participation (20%); weekly exercises you complete (30%); and your project (40%). and a short presentation of your project (10%). There are no examinations. Your score in the course will be calculated to yield your grade using this scale: A = 100-90, B = 89-80, C = 79-60, D = 59-50, F = <50. I may modify these basic grades with a + or - in special cases if I believe it is appropriate.

Basic Grading Philosophy for This Class
This class is not required for any students. Therefore, I assume that all students who have enrolled in the class have done so because they want to learn how to do data analysis. Given this, I will have little tolerance for any behavior which suggests that a student is trying to avoid learning the material. On the other hand, I encourage and try to reward behavior

which suggests that a student is attempting to enhance how quickly or thoroughly they learn the material, how to minimize the effort involved in doing an analysis correctly, and similar wholesome strategies. I will assess your understanding of the material using assignments, and each student's final write-up and presentation of their project. I will not give tests, because genuine understanding of this material is difficult to assess via a test, and because I do not want to encourage students to merely memorize material for a test.

Attendance Policy

Attendance is required at every class meeting except in the case of documented excusable absences (see the University policy on excused absences). Attendance will constitute 20 to 25% of your grade.

Policy on Collaboration and Use of Outside Resources:

Students are encouraged to work and study together during the first 2/3 of the semester, including working together on completing the weekly assignments. Additionally, there are many resources available on the internet and elsewhere, including model answers to most of the exercises in the textbook (see pp v-vi). I encourage you to use these **to the extent that they enhance your understanding of the analyses being learned.** My only requirement is that in your write-ups you must acknowledge your collaboration with other students and/or your use of these and other resources. There is never a penalty for working with other students or using additional resources so long as you acknowledge them.

However, the privilege of collaboration and use of external resources does not extend to your required individual written solution to each exercise. Each student must write up the exercises independently using their own words. You should use these write-ups to show me that you understand the analysis being performed, how to make SPSS perform the analysis, and how to interpret the output generated by SPSS. In general, the way to do this is to provide a detailed explanation of why you took the steps you chose and how you drew any interpretations you made.

Regretfully, I must punish infractions of this policy. If I find that two or more students have turned in write-ups that are copies, or which I judge to be "too similar", I will split the credit for that assignment evenly between the students involved. If I detect an answer that is too similar to the model answer on the textbook website or to those on other websites that I know of, I will at most award that student half credit.

During the last 1/3 of the semester each student will be working on their own individual data analysis project. You are welcome and encouraged to discuss your project with anybody who will sit still for it. However, you must write it up individually in your own words. Furthermore, you must acknowledge any help you got from fellow students, or anyone else, in the acknowledgment section of your final report. This principle also extends to published and online resources, which must be cited in your report and referenced in the bibliography of your report. Direct copying of published or online materials, or use of them without citation is considered plagiarism, a form of academic misconduct, and I am required by University policy give you zero credit for any assignment for which I detect it.

Weekly Assignments

You will have an assignment to do (almost) every week. The assignment will be posted on Moodle. Each assignment is explicit in what I want you to do and what I want you to submit. Most of the assignments will also include practice in writing parts of a scientific research paper.

Project

Each student will complete a project that involves analysis of a data set of their choice, applied to an anthropological problem they are interested in. Certain milestones in the completion of the project (selection of a data set, analysis results, rough draft, and final draft) will be submitted, with one or another of these due each week. The format of the paper should be scientific research paper format, which you will learn over the course of the semester. Here are some things that I will expect to see in your research paper.

1. Five part scientific format, including the following sections: introduction, materials & methods, results, discussion, and conclusions.
2. The introduction should include at least a brief literature review of other studies that have been done in the area you are working on. A minimum of 10 sources should be discussed and cited in the text of this section. These sources should be referenced in the bibliography.
3. Your paper should include a bibliography. The citation or bibliography format should be according to one of the major journals in anthropology, such as American Anthropologist, American Journal of Physical Anthropology, etc. Alternatively, you can use CSE/CBE style. Online materials are acceptable if referenced properly, and there is a large amount of advice online about how to reference online or other electronic documents.

Submission Procedures

Weekly assignments and project fragments should be submitted via Moodle . This saves me time, saves you printing costs, saves trees, and (possibly most importantly) helps me avoid losing students' work.

Assignments are due via Moodle before Tuesday at midnight during the week after they are listed in the syllabus. There will be a penalty of 20% of that assignment's score for each day (or fraction of a day) that an assignment is late.

You can expect me to grade your assignments promptly and give you feedback via a grading form and/or via comments on your assignment returned to you via your official University email address.

Other Statistical Software

As the person who has to grade your assignments, I have to standardize on one statistical software package. For many reasons, I have chosen SPSS for our standard statistical software for ANTH 402. However, there are several other commercial, shareware, and freeware statistical software packages available. In particular, I am impressed with PAST. PAST will do almost everything SPSS does, though the output isn't as pretty. PAST offers additional useful types of analysis that SPSS doesn't, such as cladistics, neighbor joining clustering, mixture analysis, and correspondence analysis. It has the best, fastest, and most flexible cluster analysis that I have ever seen. In my own research I use PAST 10 times more often than SPSS.

Advanced Anthropological
Statistics: Provisional
Schedule

This schedule is tentative and the topics might change as we go. Topics and readings will always be current on Moodle, so I have put the readings and other materials there. Thus, this schedule is a bare-bones list of topics, readings from the Handbook, and assignments. Required browsings are listed on Moodle. Assignments are due before 12:00 midnight on the Tuesday of the week after the assignment is listed unless otherwise noted on Moodle.

WEEK OF	TOPIC	HANDBOOK	ASSIGNMENT
1/25	Getting Started Intro to SPSS and the Labs **LECTURES (Notes on Moodle**) Lecture 1: Intro to the Class Lecture 2: Types of Data and Sampling	Chapter 1	Understand how the class will work Find some place to use SPSS
2/1	Basics Statistics Descriptive & Inferential Stats **LECTURES (Notes on Moodle)** Lecture 3: Descriptive Statistics Lecture 4: Inferential Statistics Lecture 5: Frequencies, Data Transformation, and Capturing Output	Chapter 2 Chapter 3	Assig 1: Descriptive & Inferential Stats
2/8	Multiple Regression **LECTURES (Notes on Moodle)** Lecture 6: Regression & Correlation Lecture 7: Working with Non-Linear Data in Regression	Chapter 4	Assig 2: Multiple Regression
2/15	Principle Component Analysis **LECTURES (Notes on Moodle)** Lecture 8: Principal Components Lecture 9: The Search for Significant Relationships	Chapter 11	Assig 3: PCA and FA
2/22	Discriminant Analysis **LECTURES (Notes on Moodle)** Lecture 10: Discriminant Analysis Lecture 11: Decisions, Decisions	Chapter 12	Assig 4: Discriminant Analysis
2/29	Cluster Analysis **LECTURES (Notes on Moodle)** Lecture 12: Cluster Analysis Lecture 13: Constructing & Reading Dendrograms		Assig 5: Cluster Analysis
3/7	Clustering PC's and DF's **LECTURES (Notes on Moodle)** Lecture 14: Clustering PC's and DF's Lecture 15: Causes of Similarity Between Things Lecture 16: Logistic Regression		Assig 6: Clustering PC & DF
3/14	ANOVA **LECTURES (Notes on Moodle)** Lecture 17: ANOVA Lecture 18: Character Coding and Missing Data	Chapter 5 & 6	Assig 7: ANOVA

WEEK OF	TOPIC	HANDBOOK	ASSIGNMENT
3/21	Beginning Your Project		Assig 8: Find a Data Set
	LECTURES (Notes on Moodle)		
	Lecture 19: Finding a Data Set		
	Lecture 20: Scientific Paper Format		
3/28	Project week 1		Assig 9: Make a Research Plan
	LECTURES (Notes on Moodle)		
	Lecture 21: Advanced Nominal Data Methods: Multidimensional scaling, classification trees, and correspondence analysis		
	Lecture 22: How to Treat Males and Females in an Analysis		
4/4	**SPRING BREAK**		
4/11	Project week 2		Assig 10: Data Analysis Results
	LECTURES (Notes on Moodle)		
	Lecture 23: Multivariate Statistics in Forensic Anthro		
	Lecture 24: The effect of Admixture on DF for Ancestry		
4/18	Project week 3		Assig 11: Preliminary Draft
	LECTURES (Notes on Moodle)		
	Lecture 25: Models of Genetic Similarity		
	Lecture 26: Intro to Some Other Statistical Software		
4/25	Project week 4		Assig 12: Revise Paper, Turn In final draft
	LECTURES (Notes on Moodle)		Due by midnight Sunday May 12
	Lecture 27: Cladistics		
	Lecture 28: Data Mining		
5/2	Project Completion: Discussion of Projects in Class		

Finals Week Tuesday May 10, 8:00am to 10:00am. Meet for project presentations by grad students