

CJ 8105
Fundamental Statistical Issues in the Analysis of Criminal Justice Data
(Statistics I)

Fall 2007

Wednesdays 1:30-5:00 (lab 4:00-5:00)

Professor: Kate Auerhahn

Office: 549 Gladfelter Hall

Phone: 204-1354
email: auerhahn@temple.edu
course web site at <http://blackboard.temple.edu>

Office Hours: Mondays 2-5 and by appointment

Welcome to Statistics!

This is a course in statistical methodology for graduate students. I will likely repeat this numerous times throughout the semester, but this is *not* a math course. Your facility with mathematics has very little to do with how you will handle statistics. Statistical thinking is about logic and relationships between (and among) things that happen to be measured quantitatively. If you can't let go of the "I hate math" thing, we're not going to get anywhere.

This being a graduate level course, it is expected that everyone will have been exposed to the basic concepts at some point before this class. This course should not be your first exposure to statistics. If you took an undergraduate statistics course sometime in the last 4 or 5 years, you should be fine. It is recommended that you re-read whatever undergraduate textbook you used (you kept it, right?) to refresh your memory if a substantial amount of time has passed. We will do a very abbreviated review to bring you back up to speed, *but I will not be spending a lot of time on the more basic issues like levels of measurement, central tendency and dispersion, and the Central Limit Theorem.* However, it is crucial that you understand these basic concepts before moving on to the more advanced applications. I have many undergraduate statistics texts in my office and would be happy to loan you one if you would like to brush up on the basics on your own.

My approach to teaching and using statistics is logic- and computer-based. Much more important than being able to hand-calculate the correct value for a

statistic is being able to actually understand what the number means. Computers will do the calculations. You need to be able to tell us what the computer output means, and why we should care (or, alternately, why we should disregard it!)

My primary aim in this course is to provide you with the tools you will need to perform and write up basic quantitative analyses. Another goal of this course is to instill in you an understanding of how to tell when those basic analyses may not be enough, and when you may need to pursue more sophisticated methodological strategies – some of which we will cover in this course, some of which will be covered in other courses. It is also important that you learn this material so that you may capably read, understand and critique the quantitative scholarly work of others in criminal justice and other social sciences.

Toward this end, this course will be very applied in nature – the bulk of your grade in the course will depend on a series of written homework assignments. I have indicated in the topic listing when an assignment will be coming, and you will generally have one week to do the assignment. All assignments will involve using SPSS to generate statistics; we will be working with the same data set (described at the end of the syllabus) for all homework assignments. We may use additional data sets for demonstration purposes in the lab from time to time.

We are scheduling a one-hour lab in addition to the regular 2.5 hour course, as approved by the department's Graduate Committee. In that lab we will run through procedures needed to complete the homework assignment for the following week, and may explore additional issues. The lab will be held in the 5th floor computer lab, right after class (so the class time will run from 1:30- 5:00 pm, with a break or two in between). In the lab, we will conduct analyses that are *just like* the ones you will need to run for your homework assignments. We will *not* be doing the homework assignments during lab periods.

Workload Expectations

Students in past years have typically reported that the time involved in this course is anywhere from 25% to 100% greater than what is required in other graduate courses. The extra effort is required because you are simultaneously learning several different things: statistical concepts, how to run and interpret statistical analyses in SPSS, and how to write up the results of these analyses. Plan to allocate more time to this course than you may be used to for other graduate courses. The graduate committee in the Criminal Justice department many years back recognized that this course is, in effect, a four credit course. The fourth credit comes because we need about an hour each week in the lab. We have not formally changed it to a four credit course due to the extra expense this

would cause for those students whose tuition is not covered by a graduate assistantship or fellowship (not to mention the fact that a single extra credit will not be particularly useful to anyone with respect to program completion requirements). BUT: plan on being in class about an hour more than the expected 2.5 per week, and plan on this course taking a lot more of your time than you thought.

Do not miss class. Really. Statistics is cumulative, and the learning process must be *active* – one implication of this is that we will move on to new topics before you become fully comfortable with the old ones. This is because the only way you get comfortable with new statistical knowledge is by *using* it. Missing class will set you back a great deal, and you will fall further behind trying to catch up. If you are going to miss class, please let me know, but please don't ask me to tell you that it is "ok" that you do so.

Grading

Homework assignments will comprise 80% of the course grade. I will drop the lowest homework grade from the average *provided* all assignments are handed in on time. If an assignment is missing or late, your homework average will be unadjusted. There are no re-writes, extensions, or extra credit in this course.

The remaining 20% of the grade will be based primarily on your participation in class (measured by your engagement in class discussions, your preparedness to discuss the assigned reading, and the like). This portion of the grade will also take into consideration (to a much lesser extent) the student's improvement and effort throughout the semester. Since attendance is required, simply attending class does not favorably influence your course grade. However, missing class more than once is likely to negatively influence this portion of your grade.

Homework

The homework assignments will involve your running some analyses in SPSS and writing up your results. These exercises are intended to serve multiple purposes. One of these is to get you comfortable with documenting data analyses and *using* statistics – the best way to learn them, in my opinion. These assignments are also writing exercises. Writing will be a large part of what all of you will be doing once you finish grad school and embark on whatever career path you choose, and the best way to improve your writing is by doing a lot of it. In the beginning of the course, I will be less concerned about your *technical* writing, and more concerned about your understanding of the statistical concepts. However, as we move through the semester and you become more comfortable with the statistical material, we will work on sharpening your

technical/statistical writing, as well as continuing to strengthen your writing generally.

You should bring two copies of your homework assignment to class on the day it is due; one to turn in, and one to make notes on as we go over the assignment. All homework assignments will be given with one week's lead time. In some cases it may be more than one week, but it will never be less than one week. Homework due dates are **not** negotiable (you can always turn in assignments early). If you turn in assignments late, 20% (or the equivalent of 2 letter grades) will be deducted right off the top for each day they are late. "Late" means anytime after 1:35 p.m. on the due date. The most challenging of these assignments may take you several hours to complete – however, you will have a period of more than 160 hours in which to complete the assignments in between receiving them and the due date. Do not miss class because you are frantically trying to complete the assignment. Time management is a critically important skill – if you are one of those people who prefers the "adrenaline rush" of leaving things until the last minute, you should *seriously* reconsider this approach with respect to statistics. Trust me. Statistics is all about precision, which is not compatible with last-minute time crunches.

All assignments should be: typed (12-point), double spaced, stapled together, including your original (not a Xerox) SPSS output. I would also be very grateful if your write-up included page numbers. Assignments should be treated as short papers, i.e. composed of complete sentences arranged in paragraphs, with flow and structure to the entire document; and perhaps most importantly, spell-checked and proofread. You should also know that abbreviations (especially "etc.") are a pet peeve of mine in course papers.

Working in groups is highly encouraged. Study groups are especially helpful in learning statistics. However, it is expected that all written work will be the sole and original product of each student. Deviations from this standard will be considered academic dishonesty and will be treated as such. Let's not go there, ok?

Blackboard Site

There is a BlackBoard site available for this course. You will need to visit it on a regular basis. If you are totally unfamiliar with BlackBoard, see me after the first class and I will bring you up to speed. I will expect to be able to communicate with the class via the BlackBoard site as well as via email, if necessary. If you use a different email than your Temple address as your primary email, then you should get that email into the BlackBoard system and/or forward your Temple

mail to the account that you do use (if you don't know how to do this, you should really just be using your Temple account).

It is essential that you utilize the Blackboard site, primarily for the *Discussion Forum*. **This is the only place I will entertain questions about the homework assignments.** I do this for two reasons: one is that lots of students usually have the same question, so it benefits everyone if those questions are asked and answered in a public forum. The second reason is that I don't want to spend my days answering the same questions over and over again in individual e-mails. I understand that grad students sometimes have this "I don't want to ask questions publicly because then everyone will know I don't get it, and I will look stupid in front of the other students" thing. Please get over that. Fast. I have news for you: when it comes to statistics, *nobody* gets it the first time. (There are a few people who do, but they are rare freaks of nature, like those people who can calculate the square root of 863 in their head). Even if you don't think you have questions, it is a good idea to visit the board and read the questions and discussions others have posted. Also - and this should go without saying - please read the whole assignment thread to see if your question has been asked and answered prior to posting.

Office Hours

I have scheduled my office hours this semester all in a block - from 2-5 pm on Monday afternoons. This time and configuration was chosen to be maximally convenient and helpful to you all. However, this time may not always work for everyone - in which case I will be happy to arrange an alternate meeting time. Please do not drop by at random times with questions about the course. I am absolutely willing (even eager) to help you in any way I can, but recognize that in addition to other teaching responsibilities, I am involved in research, writing, and a myriad of other tasks and responsibilities, and I must manage my time carefully. It may be tempting to check down the hall right when you have that "quick question" - but please resist if it's not my office hours. The reason I mention it here is that this has been a problem in the past. That's what the discussion forum is for. I check it very frequently, so you will probably not have to wait very long for an answer to your question!

Books and Readings

There is one required text for this course:

- **Knoke, David, George W. Bohrnstedt, and Alisa Potter Mee, 2002. *Statistics for Social Data Analysis (4th Edition)*. Belmont, CA: Wadsworth. ISBN 0875814484.**

This is available in the bookstore. If you can get your hands on a used third edition for less money, feel free to do so, as they are virtually identical. DO NOT buy new through an online seller, as they generally a) will not save you any money and b) will usually take a month to get it to you (because it's not exactly in great demand, most sellers don't have lots of these in stock.) If you are going to buy new, you might as well go through the campus bookstore. Statistics books are expensive; I chose carefully to select one that I think will serve you well as an all-round useful reference throughout your career. I still use my copy of this text as a reference.

- In addition to the required text, I will make copies of additional required readings available to you. Some will be made available as hard copies, and some will be posted on the BlackBoard site. These will not be especially numerous. The reading load in this course is not extensive, by graduate course standards. The citations for these readings are listed below for your reference.

Supplementary Readings (alphabetical)

Agresti, Alan, 1996. *An Introduction to Categorical Data Analysis*. New York: Wiley. (Chapters 1& 4).

Auerhahn, Kathleen, 2007. "Just Another Crime? Examining Disparity in Homicide Sentencing." *The Sociological Quarterly* 48:2:277-312.

Hagan, John, 1975. "Parameters of Criminal Prosecution: An Application of Path Analysis to a Problem of Criminal Justice." *Journal of Criminal Law and Criminology* 65:4:536-544.

Hamilton, Lawrence C., 1992. *Regression with Graphics: A Second Course in Applied Statistics*. Belmont, CA: Duxbury Press. (Chapter 1)

Hannon, Lance and Peter Knapp, 2003. "Reassessing Nonlinearity in the Urban Disadvantage/Violent Crime Relationship: An Example of Methodological Bias from Log Transformation." *Criminology* 41:4:1427-1448.

Gartner, Rosemary, 1990. "The Victims of Homicide: A Temporal and Cross-National Comparison." *American Sociological Review* 55:92-106.

Gujarati, Damodar N., 1995. *Basic Econometrics* (Third Edition). New York: McGraw-Hill. (Chapters 1, 2, 4& 6)

King, Gary 1988. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model." *American Journal of Political Science* 32:3:838-863.

Stimson, James A., 1985. "Regression in Time and Space: A Statistical Essay." *American Journal of Political Science* 29:915-47.

I also highly recommend that students purchase:

- **Norusis, Marija J. *SPSS 14.0 Guide to Data Analysis*. Upper Saddle River, NJ: Prentice Hall. ISBN 0131995286**

This is not only a very useful SPSS manual, but it's a pretty decent statistics text as well. I highly recommend it to anyone relatively inexperienced with either SPSS or statistics. I have not ordered this through the bookstore, but you can get it from amazon.com and other online booksellers quite easily. Try to get it used if you can. The 12.0 or 13.0 version will be just as useful, so if you see one of these floating around for a reasonable price (I got my copy at half.com), definitely pick it up. You may even want to share a copy among a few people.

Organization of the course:

A list of topics we will cover in the course appears below. Since I don't know how fast we will be able to go (it varies from group to group), I don't want to lock us in to a rigid schedule. I do not wish to move on until the topic has been adequately covered. That being said, it is my hope that we will spend *no more than three weeks* on the "undergraduate review" portion of the course. Anyone needing more help than that should reconsider his or her readiness to take this course, and/or see me during office hours for additional help. I am aware that you all have other courses and other responsibilities, so I will make every effort to keep you apprised of the likely schedule a few weeks in advance (when assignments will be due, etc). If this lack of a firm schedule is problematic (e.g. if you know you'll have a conflict/really busy time at some future point in the semester and you would like to plan ahead), please come talk to me about it and I'll try to give you information that will assist in planning your semester.

TOPIC LISTING

	READING
Univariate Statistics	
Levels of measurement; Distributions; Univariate Descriptive Statistics; Measures of Central Tendency and Dispersion [homework]	KBM Chapters 1 & 2
The Normal Distribution; Central Limit Theorem; Statistical Inference; Confidence Intervals	KBM Chapter 3
Bivariate Analysis	
Hypothesis Testing; t-test, Analysis of Variance, chi-square [homework]	KBM Chapters 4 & 5
Scatterplots, Correlation and Bivariate Regression [homework]	KBM Chapter 6 <i>Gujarati Chapters 1 & 2</i>
Multivariate Analysis	
Contingency Tables; statistical control [homework]	KBM, Chapter 7
Multiple regression Model [homework]	KBM Chapter 8
Assumptions of the Regression Model/ Consequences of Violation [homework]	<i>Gujarati Chapters 4 & 6</i> <i>Hamilton Chapter 1</i> <i>Hannon & Knapp (2003)</i>
Variations on the regression model: ANOVA and ANCOVA; interactions [homework]	<i>Auerhahn (2007)</i>
Categorical Dependent Variables: Maximum Likelihood Estimation; Logistic Regression [homework]	KBM Chapter 9 <i>Agresti Chapters 1 & 4</i> <i>Stimson (1985)</i> <i>Gartner (1990)</i> <i>King (1988)</i>
Special topic - Path Analysis? [homework]	KBM Chapter 11 <i>Hagan (1975)</i>
Review	

Data Set Description

We will be working with a modified version of a Bureau of Justice Statistics data collection entitled *State Court Processing Statistics: Felony Defendants in Large Urban Counties*. This data is collected yearly, and tracks felony cases filed in May until their final disposition or until 1 year has elapsed from the date of filing. This collection presents data on felony cases filed in approximately 40 of the Nation's 75 most populous counties. These 75 counties account for more than a third of the U.S. population and approximately half of all reported crimes. Data are collected on arrest charges, demographic characteristics, criminal history, pretrial release and detention, adjudication, and sentencing. More information about the data may be found at <http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/02038.xml>. The data set is available for download on the course BlackBoard site.

For a variety of reasons, I have selected a sub-set of cases from the 1998 data collection. The data set contains all cases from the states of Pennsylvania, Tennessee, and Washington (N =1205). I have also selected only 50 variables from over 100 in the original data set. Other modifications include some recoding of variables to suit our analytic purposes. This is the data set you will use for all homework assignments.