Preparation for Statistics in Sociology

August 15, 2011 – August 18, 2011
McKinney Room, Soc/Psych 329, 9:00am – 5:00pm

Instructor: Jake Fisher  
jcf26@duke.edu
Faculty Advisor: Steve Vaisey 
vaisey@soc.duke.edu

Objectives
This course provides an introduction for the two-semester graduate statistics courses by reviewing topics in research design and mathematics which underlie the practice of statistics in the social sciences. In a sense, the two semesters which follow will explain the “how” of statistics; the purpose of this course is to provide the “why”. Students will gain an understanding of the logic of variable selection, the mathematical literacy needed for graduate-level statistics, and fundamentals of statistical computing with Stata.

Course Structure
The course will begin promptly at 9:00am, and will continue with a reasonable lunch break until approximately 5:00pm each day. Morning lectures will focus on topics relevant to research design, mathematics, and introductory statistics. Afternoon labs will illustrate concepts in statistical computing with Stata.

A Note on Computers
Students should bring their own laptops. If a student does not have a laptop to use, the department can provide a laptop. Machines running Windows should be pre-loaded with an SSH client and an X window server; the department provides licenses for F-Secure and X-Win, respectively. Laptops should be used for the afternoon lab sessions only. Students should not use laptops, phones, tablets, etc. during the morning lectures.

Class Schedule

Monday, August 15

Morning: Research design, calculus
- Research design: types of variables, measurement, logic of causal order
- Calculus overview: limits, derivatives, integrals

Afternoon: Introducing Stata
- Opening Stata: forwarding X windowing system, SSH
- Using the server: mapping home directory, absolute and relative paths
- Opening datasets in Stata
- Stat/transfer

Acknowledgements: This course draws inspiration and material from similar offerings in the Duke Political Science department, taught by Sunshine Hillygus, Jacob Montgomery, and Chris DeSante.
Assignment:  ASR articles
Read two articles from a recent issue of The American Sociological Review, distributed in class, with an eye towards variables used, measurement, and causal pathways. Everyone will present one of these articles on Tuesday.

Tuesday, August 16

Morning:  Research design (continued), matrix algebra, probability
Discussion of assigned articles
Scalars, vectors, matrices
Matrix operations
Probability: sets

Afternoon:  Variables in Stata
Good programming: logging, commenting, do-files
Variables: labels, formats, missing values
Logical operators
If-then statements
Recoding variables
Subsetting data by variables and by observations: keep, drop, preserve, restore

Wednesday, August 17

Morning:  Probability (continued), statistics
Distributions, marginal and joint
Expected values
Univariate statistics
Hypothesis testing

Afternoon:  Tables in Stata
Advanced recoding: loops, expressions with gen
Cross-tabulations
Macros
Replicate Vaisey

Thursday, August 18

Morning:  Statistics exercises
Practice makes perfect!
Other topics in statistics as desired, time permitting

Afternoon:  Advanced topics in Stata
By-groups and sort
t-tests
Loading packages: publication-ready tables with estout, outreg
Linear regression
Using saved values (from functions)
Graphing
References


