56:824:702:01 Quantitative Methods I
Fall 2017, Mondays, 6:00pm to 8:50pm, BSB 133

Syllabus, draft June 27, 2018

Professor: Paul A. Jargowsky; 856-225-2729; paul.jargowsky@rutgers.edu
Office Hours: Cooper 321 (first floor), Mondays 4:00-6:00pm and by appointment.

> Statistics are not magical. Nor are they always true – or always false. Nor need they be incomprehensible. Adopting a critical approach offers an effective way of responding to the numbers we are sure to encounter. Being critical requires more thought, but failing to adopt a critical mind-set makes us powerless to evaluate what others tell us. – Joel Best, "Telling the Truth about Damned Lies and Statistics," *The Chronicle of Higher Education*, May 4, 2001.

Overview. This course is designed to prepare students for advanced quantitative methodology courses required of doctoral students. The course begins by reviewing descriptive statistics and data presentation techniques. In preparation for the study of inferential statistics, the next section of the course covers the basics of probability. A solid grounding in probability is necessary to understand how and why statistical techniques work. Building on that foundation, the heart of the course is a rigorous introduction to statistical inference: sampling theory, confidence intervals, and hypothesis tests. The final section of the course is an introduction to regression analysis, with an emphasis on interpretation of regression results, using examples from recent research. This course is part of a two semester sequence; the second semester is Quantitative Methods II, which is a more advanced and detailed treatment of regression analysis and related topics.

Learning Objectives. The main objective of the course is that students will develop a conceptual and practical understanding of how to learn from social science data; specifically:

- Students will learn how to calculate and interpret descriptive statistics;
- Students will have a grounding in probability and understand how to work with random variables, and univariate probability distributions (including the normal, binomial and Poisson distributions);
- Students will be able to calculate and interpret correlation coefficients;
- Students will understand the theory behind inferential statistics that supports drawing conclusions about a population from a sample;
- Students will be able to calculate confidence intervals and conduct basic hypothesis tests for means, proportions, and variances, including large and small samples;
- Students will understand Type I and Type II errors and will be able to calculate the power of the test;
- Students will be able to read and interpret basic bivariate and multiple Ordinary Least Squares regressions.

Prerequisites. The topics discussed in this course are inherently mathematical. Students taking this course should be very comfortable with college algebra. Some calculus will be used, but no knowledge of calculus is assumed or required.

Textbooks. The primary textbook for this course is an open source, online textbook that is freely available online:

> David M. Diez, Christopher D. Barr, and Mine Çetinkaya-Rundel, OpenIntro Statistics, 3rd Ed. Available from http://www.openintro.org.

Note: I do not assign the readings from the text in the order in which they appear in the book. Thus, *you need to pay attention to the specific sections that are assigned for each class*. Readings are indicated in the syllabus by chapter and section number. If "Section 1.2" is listed, that means all of Section 1.2, including any subparts like 1.2.1, 1.2.2, 1.2.3, etc. If I only want you to read particular subparts, I will list them specifically (1.2.x) and you can safely skip the other parts. I will make the text available on Sakai. If you prefer a hard copy, you may order a printed copy from Amazon.com for a small fee. Other reading materials will be made available via Sakai.

Computing. This is a course about concepts, not software. Nevertheless, we will use software at times to reduce the computation burden. *Stata*, version 15, is the official software for the class, but any recent version of Stata will work nearly as well. *You do not need to buy Stata.* It is available on the Rutgers on-line system (apps.rutgers.edu) and in the Robeson Computer Lab. However, if you wish to use Stata on your own devices, a special discounted version will be available to students through the "grad plan." You can get a six-month license for $45, though you may want to consider the annual or perpetual license, since Stata will be used in future courses. To order, please contact StataCorp directly: Phone: 800-782-8272 (Monday through Friday 8:00 to 5:00 Central Time) or order online at https://www.stata.com/order/new/edu/gradplans/student-pricing/. Be sure to ask for the "grad plan." If ordering online, use your Rutgers email address to verify affiliation with the university.

Miscellaneous. A *calculator* is a necessity. It does not have to be fancy and it does need graphics or programming capabilities, but it is useful to have the following functions: $\ln(x)$, e^x, $x!$, and y^x (logarithms, exponents, factorials, and powers). Usually any calculator described as "scientific" will have these functions. Good choices are available for as little as $10. Even the built-in iPhone calculator will suffice; access the scientific functions by turning the phone sideways.

Grading and Requirements:

Problem Sets: The only sure way to learn the material presented in this course is to work on problems that reinforce the readings and lectures. Thus, there will be a short problem set due almost every week (see the schedule below for exact due dates). Except in unusual

circumstances, *late problem sets will not be accepted* because the correct timing of the work is important in the learning process and because the answers will be discussed in class. However, *the lowest two problem set grades will be dropped*, allowing the student some flexibility and margin for error.

Tests: There will a midterm examination on October 29. The final exam, scheduled for Dec. 17, is comprehensive. The tests are *open-book, open-note*. A calculator is a necessity, hopefully one with which you are familiar. *Laptop computers are not permitted during the test.* Mark your calendar now, because in fairness to other students, I cannot create make-up tests or reschedule tests for any one person.

Grades will be assigned as follows: Letter grades will be determined based on the overall course average, rounded to the nearest whole number. Only certain letter grades are available for graduate course work at Rutgers-Camden. The translation of a numeric grade to a letter grade will be done as follows:

A	90-100
B+	85-89
B	80-84
C+	75-79
C	70-74
F	0-69

In determining the course average, assignments will be weighted as follows:

Problem sets (lowest 2 dropped)	40 percent
Midterm exam	20 percent
Short paper assignment	10 percent
Final exam	30 percent

At any point in the semester, you can see your current course grade in Sakai based on the graded work up to that point. You can also figure out what grade you need on the final exam to get any specific final grade.

Incompletes: Generally speaking, the material in this course is best learned as a single unit. I will grant incompletes only in cases where a substantial change in life circumstances occurs that is beyond the control of the student, and then only with appropriate documentation.

Ground Rules for the Course:

Attendance. Attendance is entirely optional. Having said that, you are responsible for everything covered in class whether it was covered in the readings or not. You are also responsible for any announcements made in class. For most people, attendance is a necessary condition for learning the material. The PowerPoint slides are not a substitute for attending the class, because the slides are not self-contained – they are props to give structure to my lectures

and our class discussions. If you need to miss a class, be sure to get notes and a recap from a classmate.

Study Groups. You are encouraged to form and participate in a regular study group. Many students over the years have found the study groups to be very helpful. Study groups are permitted and encouraged to work on the problem sets together. However, *each individual student should write up his or her own answer to hand in, based on his or her own understanding of the material. Do not hand in a copy of another person's problem set, even a member of your own group.* Writing up your own answer helps you to internalize the group discussions and is a crucial step in the learning process.

Academic Integrity. Violations of academic integrity include cheating on tests or handing in assignments that do not reflect your own work and/or the work of a study group in which you *actively* participated. *I have a policy of zero tolerance for cheating.* Violations will be referred to the appropriate university authorities.

Course Schedule. The course schedule below is for illustration only. Always consult the online version for the latest dates, readings, and assignments.

Schedule, Readings, and Assignments **Draft, June 27, 2018** **Consult online version on Sakai for latest version** *Note: the links below work best in the Chrome or Firefox browsers* *The Sakai help desk suggests that you avoid Internet Explorer when using Sakai.*	
Date	**Title of Class** *Topics* *Readings* *Assignments Due*

Sep. 10	**Descriptive Statistics**
	
	Topics
	
	• Types of Data
	• Measures of Central Tendency
	• Measures of Variability, skewness, etc.
	• Correlation between variables
	• Graphic Techniques (Histograms, Box Plots, etc.)
	
	Readings
	
	• Jargowsky and Yang, *Descriptive and Inferential Statistics*, Section I, pp 1-16 only.
	• Joel Best, "Telling the Truth about Damned Lies and Statistics," *The Chronical of Higher Education*, May 4, 2001.
	• OpenIntro Text
	  ○ Section 1.2, "Data Basics"
	  ○ Section 1.3, "Overview of data collection principles"
	  ○ Section 1.6, "Examining Numerical Data"
	  ○ Section 1.7.1, "Contingency Tables and Bar Plots"
	  ○ Section 1.7.2, "Row and Column Proportions"
	
	Introduction to Stata Interface
	
	• https://www.youtube.com/watch?v=L8iIj_8lhRc

Sep. 17	**Effective Presentation of Data** *Topics* • Identifying the informative contrast • Effective communication with tables and graphs *Readings* • Gene Zelazny, *Say it with Charts*, Introduction • Gene Zelazny, *Say it with Charts*, Section 1: Choosing Charts • Tufte, *The Decision to Launch the Space Shuttle Challenger* • (optional) Watch the Challenger disaster https://www.youtube.com/watch?v=AfnvFnzs91s • Tufte, "The Cognitive Style of Powerpoint" *Assignments Due* • *Problem Set #1*

Sep. 24	**Basics of Probability** *Topics* • Simple and Compound Events • Rules of Probability • Bayes' Theorem • Random Variables • Expected Value and Variance of Random Variables *Readings* • Open Intro Text, Chapter 2, "Probability" • Gina Kolata, "Mammogram Talks Prove Indefinite," *The New York Times* • (optional) Stock and Watson, Chapter 2: Review of Probability, pp. 14-35.

Oct. 1	**Probability Distributions**
	
	Topics
	
	• Discrete probability distributions
	• Continuous probability functions
	
	Readings
	
	• OpenIntro Text:
	  ○ Section 3.1, "Normal Distribution"
	  ○ Section 3.4, "Binomial Distribution"
	  ○ Section 3.5.2, "Poisson Distribution"
	
	Assignments Due
	
	• Problem Set #2

Oct. 8	**Sampling Theory and Statistical Inference**
	
	• Basic principles of sampling
	• Sampling error vs. bias
	• The Central Limit Theorem
	• Confidence intervals for means and proportions
	
	Readings
	
	• Jargowsky and Yang, *Descriptive and Inferential Statistics*, Sections II and III, pp 16-26.
	• Lienesch et al., "How Much Can we Trust the Polls?" Huffington Post, Oct. 3, 2014.
	• OpenIntro Text:
	  ○ Section 1.3, "Overview of data collection principles"
	  ○ Introduction to Chapter 4
	  ○ Section 4.1, "Variability in estimates"
	  ○ Section 4.2, "Confidence Intervals"
	  ○ Section 4.5.1, "Confidence Intervals for nearly normal point estimates."
	  ○ Section 6.1.1, "Identifying when the sample proportion is nearly normal"
	  ○ Section 6.1.2, "Confidence intervals for a proportion"
	  ○ Section 6.2.1, "Sample distribution of the difference of two proportions"

Oct. 15	**Introduction to Hypothesis Testing** *Topics significance* • Confidence intervals for means and proportions • Introduction to hypothesis testing *Readings* • OpenIntro Text: ○ Section 4.3, "Hypothesis Testing" ○ Section 4.4, "Examining the Central Limit Theorem" ○ Section 4.5.2, "Hypothesis testing for nearly normal point estimates" ○ Section 4.5.3, "Non-normal point estimates" ○ Section 4.5.4, "When to retreat" ○ Section 4.5.4, "Statistical significance versus practical significance" ○ Section 6.1.3, "Hypothesis testing for a proportion" • Dan Myers, "Along Polluted Rio Grande, Rise in Birth Defects Brings Fear," *The Boston Globe*, September 7, 1993, p. 3. *Assignments Due* • Problem Set #3
Oct. 22	**Hypothesis Testing II: Small Samples and Differences between Samples** *Topics* • Student's t distribution • Hypothesis tests for small samples • Hypothesis tests for differences between samples *Readings* • OpenIntro Text: ○ Section 5.1, "One-sample means with the t-distribution" ○ Section 5.2, "Paired data" ○ Section 5.3, "Difference of two means" ○ Section 6.2.3, "Hypothesis tests for p1 − p2" • "Deprogramming Heaven's Gate," *The New Yorker*, April 14, 1997, pp. 31. *Assignments Due* • Problem Set #4

Oct. 29	**Midterm Examination** Open Note, Open Book *Bring a calculator!*

Nov. 5	**Hypothesis Testing III: Statistical Power, Chi-Squared Tests**
	
	Topics
	
	• The power of the test
	• Tests of Goodness of Fit
	• Tests of Association between Categorical Variables
	
	Readings
	
	• OpenIntro Text:
	  ○ Section 5.4, "Power Calculations"
	  ○ Section 6.3, "Testing for goodness of fit using chi-square"
	  ○ Section 6.4, "Testing for independence in two-way tables"
	• Stata, Conceptual Intro to Power (YouTube)
	• Stata, Tour of Power Calculations (YouTube)
	• Friedman, E. "Pet Ownership and Coronary Heart Disease Survival." *Circulation*, Vol. 168, pp. 57-58 (1978).

Nov. 12	**Hypothesis Testing VI: Tests of Variances, Bootstrapping** *Topics* • Sampling distributions for a variance • Hypothesis tests regarding a variance • Comparing 2 variances • ANOVA *Readings* • Mendanhall, Beaver and Beaver, Sections 10.6-10.8 • Open Intro Text: ○ Sections 5.5 (intro), 5.5.1-5.5.3, "Comparing many means with ANOVA" • Stata YouTube video tutorial on One-Way ANOVA • Stata YouTube video tutorial on Two-Way ANOVA • Aberson, Chris (2002). "Interpreting null results: Improving presentation and conclusions with confidence intervals." *Journal of Articles in Support of the Null Hypothesis, 1, 36-42.* *Assignments Due* • Problem Set #5

Nov. 19	**Hypothesis Testing V: When the Sampling Distribution is Unknown** *Topics* • Bootstrapping • Non-parametric tests *Readings* • Karl Wuensh, "Nonparametric Statistics." • Fruend, Mathematical Statistics, Sections 16.1-16.4. • Robert Stine. 1989. "An Introduction to Bootstrap Methods," *Sociological Methods and Research*, pp. 243-250. Note: you only need to read the first 8 pages! *Assignments Due* • Problem Set #6
Dec. 3	**Bivariate Regression** *Topics* • The population regression function • The sample regression function • Obtaining the estimates • Testing the hypothesis that X affects Y *Readings* • Tufte, Chap. 3, pp. 65-91. • (optional) Tufte, Chap 3, pp. 91-end

Dec. 10	**Multiple Regression** *Topics* • Isolating the Effect of X on Y • Left Out Variable Bias • The true meaning of multiple regression • Hypothesis testing • Goodness of Fit *Readings* • Tufte, Chapter 4. • Devaney, Barbara, Linda Bileimer, and Jennifer Schore. 1992. "Medicaid Costs and Birth Outcomes: The Effect of Prenatal WIC Participation and the Use of Prenatal Care." *Journal of Policy Analysis and Management* 11: 573-92. *Assignments Due* • Short Paper Assignment Due
Dec. 17	***** ***Final Exam*** ***** *Open book, open note.* *Bring a calculator!*