

POLS 706 PSYCH 650/790 : Spring 2015

Paul E. Johnson
Professor, Political Science
Office Hours: T9:30-11:00AM,R 3-5PM & by appt.

pauljohn@ku.edu
Acting Director, Center for Research Methods and Data Analysis
470 Watson Library

CRMDA: (785)864-3353

GTAs:

- Leslie Shaw <leslie.shaw@ku.edu> Fraser 10 To Be Determined
- Ben Rogers <bjroger@ku.edu> Blake Hall 311 TF 1:30-3:00pm

KU BlackBoard: <http://courseware.ku.edu>. Has announcements, readings, and exercises.

Course Webpage <http://pj.freefaculty.org/stat>

1. Lectures & Guides, topically organized: <http://pj.freefaculty.org/guides>.
2. There's a new lecture map on <http://pj.freefaculty.org/stat/lectures.html>.

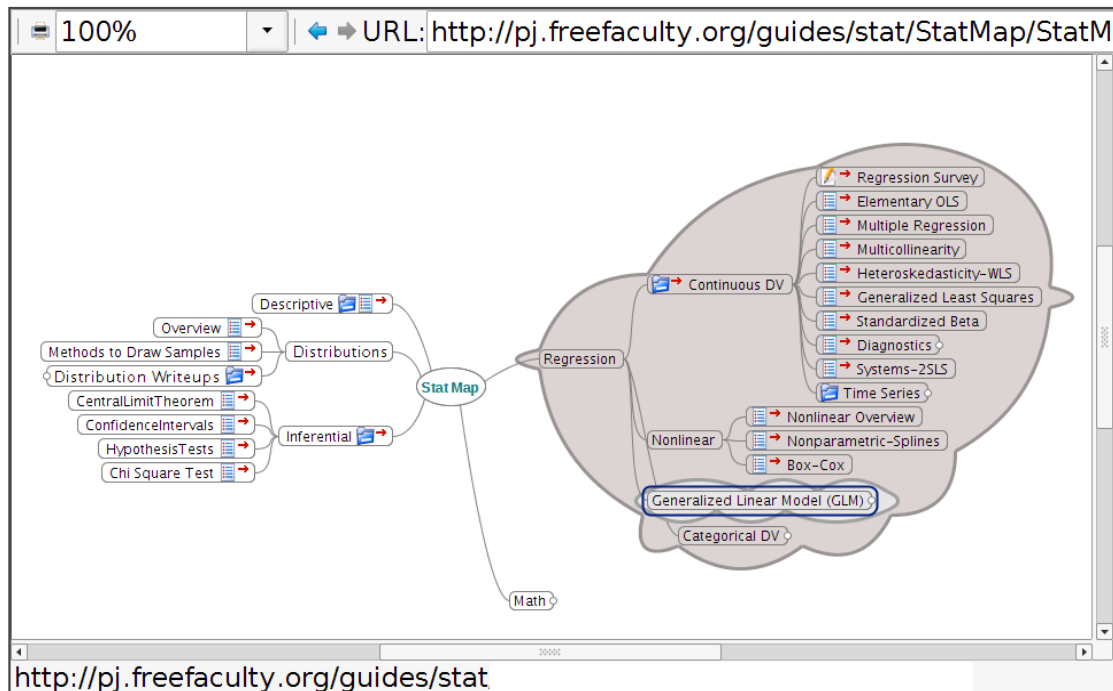
CAUTION: I'll be updating the lectures as we go through the semester, usually the night before class.

3. R support page. <http://pj.freefaculty.org/R>.

Here's a Mind Map representation of the stat section.

PJ's Re-organized Stat Writeups

<http://pj.freefaculty.org/guides>



Textbooks

In Blackboard, you should find articles/chapters of many books

Verzani, John. 2005. *Using R for Introductory Statistics*. Chapman & Hall/CRC.

John Fox. 2008. *Applied Regression Analysis and Generalized Linear Models*. Beverly Hills: Sage.

Johnson, Paul. *Stuff Worth Knowing*. Newest edition always available in the KU Blackboard system.

Course Topics: distributions, inference, regression analysis

Prerequisites: One semester of statistics and/or research design. Algebra is vital and will be used regularly. Calculus is not required, but helpful in understanding derivations.

Grades

Tests: Midterm Exam: 2 x 25%

Final Exam: 30%

Homework Exercises: 20%

This Course is Half Flipped

“Flip the Classroom” = record some lectures and ask students to watch them before class, then use class time for exciting, engaging exercises. *I am not entirely convinced this is a good idea*, but I think it is interesting. In 2014, I started recording some lectures and putting them online.

- 2 places to Download. Blackboard and also in my personal webpage.
- OGV Format. I used the “ogv” format because it is viewable inside modern Web browsers.

Lectures for all Wednesdays, and some Mondays will be available.

Lab

Lab sessions will begin the second week of the semester.

For the people that enroll in this as a 4 credit course, attendance at lab sessions is **REQUIRED**.

For others, attendance at lab is optional and **STRONGLY RECOMMENDED**.

We will have two times for lab participation per week.

1. Wednesday 4:30pm-5:15pm. Hopefully, we will get permission to use 2 rooms, so as to make small, pleasant groups.
2. Thursday or Friday (TBD): Another mutually agreed upon time that we will choose on the first day of class.
3. Monday (TBD): Homework Study Hall!

Please Use the Blackboard Forum! Don't Ask For Help In Private Emails.

We have a “forum” (or “bulletin board”). Please ask questions there, so everybody gets the same information. If you are asking an R question, please include this information.

1. Output from R command “sessionInfo()”.
2. R code to reproduce the problem you are seeing.
3. If you are asking about an error message, paste in the R output so we can see the exact wording.

Computer Tips

1. Keep your exercises projects organized into separate folders. BE HIERARCHICAL.

```
pols706/  
  assign_1  
  assign_2  
  pj-lectures
```

Keep It Separated so that, all of the data and code for a particular thing stays together. This makes it reasonable for me to say to you, “send me a zip file for that assignment” and I’ll run your code.

I’m keeping project folders subdivided into additional compartments “R”, “data”, and so-forth. Can show you what I mean

2. Never put blank spaces in file names or directory names. Also don’t use symbols like ! * & or []
3. Join email lists like “r-help”, (<http://www.r-project.org>), semnet (<http://www2.gsu.edu/~mkteer/semnet.html>), etc. I suggest you get an account on Gmail for those lists (and learn how to use labels and filters on Gmail).
4. Learn to make professional looking tables and graphs.
5. Consider learning to 1) backup your files and 2) use a version management system (Git or Subversion, for example)

Conduct

Students are expected to comply with University codes, policies, laws, and guidelines. Information on academic misconduct is available at the following address: <http://www.studenthandbook.ku.edu/codes.shtml#Academic%20Misconduct>. The work that students submit must be their own work. We recognize the fact that people do study together and exchange ideas, but, at the end of the day, each student is responsible for doing her/his own work.

Schedule of Readings

1. Jan 21. Welcome, Orientation.
 - (a) Make sure you can run R. The trick is in getting a Good Editor and developing good habits to keep files organized. People who aspire to be R programmers should install Emacs (for Windows & Mac, see Vincent Goulet’s website). R novices may prefer R Studio.
 - (b) Document preparation. I use L^AT_EX (guides <http://pj.freefaculty.org/latex>). If you use something else, GET GOOD AT IT. Learn to write math formulae in your program. Learn how to easily insert graphs and update them.
 - (c) Windows Tip: Turn off “hide extensions of known file types” in Windows Explorer View options.
 - (d) Math Refresher: If you have forgotten college algebra, please consider some Emergency Math Review. See *Stuff Worth Knowing*, “Summation Signs” & “Curves in Theory”. Several of us prepared slides on these things for the Summer Institute in 2014:
 - i. Functions <http://crmda.dept.ku.edu/SI2014/01.01-functions/presentation/functions-1.pdf>
 - ii. Math Notation http://crmda.dept.ku.edu/SI2014/01.02-symbols/presentation/Deciphering_Math_Notation.pdf
 - iii. Calculus <http://crmda.dept.ku.edu/SI2014/02.01-calculus/presentation/calculus-1.pdf>

- iv. Matrix Algebra <http://crmda.dept.ku.edu/SI2014/02.02-matrices/presentation/matrices-1.pdf>

stat/Welcome Slides: Welcome.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/Welcome/Welcome.pdf>

stat/Regression/Overview Slides: Regression-Overview-Lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/Overview>

2. Jan 26. Describing One Variable (primarily numeric variables)



Before Class Read: Verzani, Ch. 1 & 2, Fox, Chapter 3.1

Priorities: histogram, kernel density, mean, median, mode, standard deviation, variance, scaling

Slides: CentralTendencyAndDispersion.pdf

URL: <http://pj.freefaculty.org/guides/stat/Descriptive/CentralTendencyAndDispersion/>

3. Jan 28. Describing Two Variables: scatterplot, barplot, boxplot, correlation



Before Class Read: 1) Verzani, Ch. 3

2) Fox, Chapter 3.2, 3.3



Before Class View: "ScatterBoxBarPlots-1-lecture.ogv"

Slides: "ScatterBoxBarPlots-1-lecture.pdf"

URL: <http://pj.freefaculty.org/guides/stat/Descriptive/ScatterBoxBarPlots>.

4. Feb. 2. Statistical Distributions



Before Class Read: 1) Verzani, Chapter 5

2) Paul E. Johnson, "Distribution Overview: Probability by the Seat of the Pants," <http://pj.freefaculty.org/guides/stat/Distributions/DistributionOverview/DistributionReview.pdf>. *An essay I have created just for this purpose!* Same as last chapter in *Stuff Worth Knowing*.

3) Skim this: *Regress+ Compendium of Statistical Distributions* I put a copy on the Blackboard http://www.causascientia.org/math_stat/Dists/Compendium.pdf



Before Class View: "DistributionReview-1-lecture-part1.ogv"

Slides: DistributionReview-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Distributions/DistributionOverview>

I'm Unsure how much time there will be, but this class, or during the next one, I'll show you this piece.

Slides: rRandomVariables.pdf

URL: <http://pj.freefaculty.org/guides/Rcourse/rRandomVariables>

5. Feb. 4. Statistical Distributions (continued): Using R to Explore Distributions:



Before Class Read: Verzani, Chapter 6



Before Class View: “DistributionReview-1-lecture-parts2-3.ogv” (2nd half of previous lecture)

Slides: Same as last class

6. Feb 9. Central Limit Theorem and Sampling Distributions



Before Class Read: Bowen & Weisberg, *An Introduction to Data Analysis*, Ch. 10 “Statistical Inference” (available in the Blackboard)

Slides: CLT-lecture1.pdf

URL: <http://pj.freefaculty.org/guides/stat/Inferential/CentrallimitTheorem>

7. Feb 11. Confidence Intervals



Before Class Read: Verzani, Ch. 7



Before Class View: “ConfIntervals-lecture.ogv”

Slides: “ConfIntervals-lecture.pdf”

URL: <http://pj.freefaculty.org/guides/stat/Inferential/ConfIntervals>

8. Feb 16. Hypothesis Testing



Before Class Read: Verzani, Ch. 8

Slides: SignifTests-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Inferential/HypoTesting>

9. Feb 18. Regression: One Input Variable



Before Class Read: 1) Verzani, Ch. 3.4, and Ch. 10;

2) Fox, Ch 5.1

3) Woolridge, *Introductory Econometrics*, Chapter 2. This is a VERY EXCELLENT regression introduction.



Before Class View: Regression-1-lecture.ogv

URL: <http://pj.freefaculty.org/guides/stat/Regression/ElementaryOLS>

10. Feb 23. Regression: Hypo testing: t tests, F test, Confidence Intervals



Before Class Read: 1) Fox, Ch. 6.1

Slides: Regression-2-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/ElementaryOLS>

11. Feb 25. More Worked Examples

Examples! In this folder,

<http://pj.freefaculty.org/guides/stat/Regression/ElementaryOLS>

I so far have 3 full worked examples of elementary regression

Look for files named Regression-Example-1, and so forth.



Video: Not prepared yet, but will do a couple soon

12. Mar 2. Regression: Multiple Inputs



Before Class Read: 1) Verzani, Ch. 11
2) Wooldridge, Introductory Econometrics, Ch 3 (Blackboard).
3) Fox, Ch 5.2-end, Ch 6.2-end

Slides: Regression-MultipleInputs-lecture-1.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/MultipleRegression>

13. Mar 4. Regression II: Variable Selection, Variable Importance



Before Class Read: Cohen, Ch. 5.5



Before Class View: Regression-MultipleInputs-lecture-2.ogv

Slides: Regression-MultipleInputs-lecture-2.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/MultipleRegression>

14. Mar 9. Test 1. This will include ONLY material up to class session 12. Does not include Regression with Multiple Inputs

- (a) Suppose the test itself requires 90 minutes.
- (b) Then we rest for 10 minutes.
- (c) Categorical Predictors. From Model “frame” to “Design matrix”

Spring Break Mar 16-Mar 22

15. Mar 11. Regression Assumptions and Plots For Checking Them



Before Class Read: Cohen, Ch. 4
This foreshadows Fox, Ch 12.3, “Component Plus Residual” plots



Before Class View: Regression-MultipleInputs-lecture-3.ogv

Slides: Regression-MultipleInputs-lecture-3.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/MultipleRegression>

16. Mar 23. Topic: Multicollinearity



Before Class Read: Fox Ch. 13

Slides: Multicollinearity-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/Multicollinearity>

17. Mar 25. Standardized Regression & R^2 are over-rated



Before Class Read: King, Gary. (1986). "How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science*, 30(3), 666-687.



Before Class View: Standardized-1-lecture.ogv will be prepared, I promise!

Slides: Standardized-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/StandardizedBeta>

18. Mar 30. Nonlinear Regression I



Before Class Read: Fox, Ch. 12.3, Fox, Ch 17

Slides: Nonlinear-1-Overview-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression-Nonlinear/Nonlinear-Overview/>

19. Apr 1. Nonlinear Regression, II



Before Class Read: Fox, Ch 18



Before Class View: Nonparametric-1-lecture.ogv

Slides: Nonparametric-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression-Nonlinear/Nonparametric-Loess-Splines>

20. Apr 6. Outliers and Influence Diagnostics (The Hat Matrix)



Before Class Read: 1) Fox Ch. 11

2) Cohen Ch. 10.1-10.5

Slides: RegDiagnostics-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/RegressionDiagnostics>

21. Apr 8. Interactions: "mean centering" does not ameliorate multicollinearity. Period.



Before Class Read: 1) "rockchalk package Vignette". This is distributed with the package, you might as well know it exists.

2) Please read this for background: Cohen, Ch. 7



Before Class View: Interaction-Continuous-1-lecture.ogv

Slides: Interaction-Continuous-1-lecture.pdf

Lecture: /stat/Regression-Nonlinear/Interaction-Continuous <http://pj.freefaculty.org/guides/stat/Regression-Nonlinear/Interaction-Continuous/>

Additional Reading: More on why Mean Centering does not help with Multicollinearity

Echambadi, R., & Hess, J. D. (2007). Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models. *Marketing Science*, 26(3), 438-445.

Kromrey, J. D., & Foster-Johnson, L. (1998). Mean Centering in Moderated Multiple Regression: Much Ado about Nothing. *Educational and Psychological Measurement*, 58(1), 42 -67.

22. Apr 13. Categorical Predictors



Before Class Read: Fox, Ch. 7.

Most of my lecture is an effort to understand terminology in this chapter, which you might want to glance through: Cohen, Ch. 8

Lecture(after test): /stat/Regression/CategoricalPredictors/lecture-1

<http://pj.freefaculty.org/guides/stat/Regression/CategoricalPredictors/>

23. Apr 15. Interactions: Categorical with Continuous



Before Class Read: 1) Review Fox, Ch. 7.
2) Cohen, 9.3



Before Class View: Interaction-Categorical-1-lecture.ogv

Lecture: /stat/Regression-Nonlinear/Interaction-Categorical <http://pj.freefaculty.org/guides/stat/Regression-Nonlinear/Interaction-Categorical/Interaction-Categorical-1-lecture.pdf>

24. Apr 20. Heteroskedasticity: An excuse to talk about MultiLevel Models



Before Class Read: Fox, Ch. 12.2

Slides: Heteroskedasticity-WLS-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression/Heteroskedasticity-WLS/>

Note: I am removing “time series autocorrelation” from this lecture material.

25. Apr 22. TBA. I'd like to make up a lecture about the Missing Data problem.

26. Apr 27. Midterm Exam #2

27. Apr 29. Logistic Regression



Before Class Read: Verzani, Ch. 12; Cohen, 13.1-13.2



Before Class View: Regression-1-lecture.ogv

This guide on Logit and Probit models was perfected over years and years <http://pj.freefaculty.org/guides/stat/Regression-Categorical/LogitProbit/LogitProbit-1-guide.pdf>. The lecture slideshow is not quite so well evolved...

Slides: LogitProbit-1-lecture.pdf

URL: <http://pj.freefaculty.org/guides/stat/Regression-Categorical/LogitProbit/>

28. May 4. Optional Topic: Logit II



Before Class Read: Same

URL: <http://pj.freefaculty.org/guides/stat/Regression-Categorical/LogitProbit-WorkedExample>

29. May 6. Review

Final Exam: Thursday, May 14, 4:30-7:00pm (Assumes I've read this correctly: <http://registrar.ku.edu/spring-2015-final-exam-schedules>).