Applied Quantitative Methods for the Social Sciences I

Program: Master in Social Sciences, Spring 2023
Room: 18.1.A04
Time: Thursdays, 10:00–13:00

Contact Information

Instructors:	Patrick Kraft, PhD	Guillermo Kreiman, PhD
Office:	18.2.A.19	18.2.A.36
Email:	patrickwilli.kraft@uc3m.es	gkreiman@clio.uc3m.es
Office Hours:	Thursdays, 14:30–16:30, or by appointment.	By appointment.

I   Course Description

This class is a second graduate course for political methodology that builds explicitly on the foundation laid in AQMSS-I. In that class, you learned the nuts and bolts behind statistical inference as well as the statistical software R. In this course, we extend those tools to cover various linear and non-linear models. Therefore, the main goal for this semester is not only to become a proficient *consumer* of quantitative research, but also to set the groundwork to pursue your own projects using advanced statistical modeling techniques. The topics that we will intensively discuss include:

1. Ordinary Least Squared (OLS)
2. OLS diagnostics
3. Maximum Likelihood Estimation (MLE)
4. Generalized linear models
5. Model visualization and interpretation
6. Causal inference.

II   Textbook and Additional Material

Required

There is one required textbook for the course:
Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and other stories*. Cambridge University Press.
You can download it for free at https://users.aalto.fi/~ave/ROS.pdf

Any additional required readings will be available electronically on Aula Global.

Furthermore, I recommend the following resources that are available online:
- Tidyverse Cookbook (Grolemund, 2020): https://rstudio-education.github.io/tidyverse-cookbook/
- R Graphics Cookbook (Chang, 2018): https://r-graphics.org/
- R for Data Science (Wickham and Grolemund, 2016): https://r4ds.had.co.nz
- Data Visualization (Healy, 2018): https://socviz.co
- Fundamentals of Data Visualization (Wilke, 2019): https://clauswilke.com/dataviz/

There are countless other resources available online, but I want to highlight a great set of YouTube videos in case you want to learn more about specific topics covered in our course. You'll find links to these videos on Aula Global as well:
- Gary King's lecture videos on quantitative social science methods:
  https://www.youtube.com/channel/UCtrwX29xpuWc9y0-0PKrHSQ/playlists
  https://projects.iq.harvard.edu/gov2001

We will be using RStudio for the programming portion of the course. You can get started by installing R and RStudio on your computer. Next, you can work through RStudio's primers, a set of interactive tutorials that will help you familiarize yourself with basic programming concepts and R.

Additional Readings (Optional)

King, Gary. 1998. *Unifying political methodology: The likelihood theory of statistical inference*. University of Michigan Press

Fox, John. 2015. *Applied regression analysis and generalized linear models*. 3 ed. Sage Publications (older edition is fine)

Fox, John, and Sanford Weisberg. 2018. *An R companion to applied regression*. 3 ed. Sage Publications

Wooldridge, Jeffrey M. 2013. *Introductory econometrics: a modern approach*. Cengage Learning

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press

Long, J Scott. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks: Sage Publications (newer version using Stata is also available).

III Schedule

Date	Lecture	Lab	Readings	Assignments
02/02	Introduction	Intro to R and the tidyverse I	ch. 1-2	
09/02	Statistical Inference	Intro to R and the tidyverse II	ch. 3-5	
16/02	Linear Regression Review I	Data Visualization I (ggplot)	ch. 6-8	
23/02	Linear Regression Review II	Data Visualization II (ggplot)	Aula Global	PS 1 (Coding)
02/03	Bayesian Inference	Linear Regression in R	ch. 9-10	
09/03	Model Evaluation	R Case Studies	ch. 11	PS 2 (Theory)
16/03	Transformations	R Case Studies	ch. 12	
23/03	Logistic Regression I	R Case Studies	ch. 13	PS 3 (Coding)
30/03	Logistic Regression II	R Case Studies	ch. 14	
06/04	**Easter Break – No Class**			
13/04	Generalized Linear Models I	R Case Studies	ch. 15	PS 4 (Theory)
20/04	Generalized Linear Models II	R Case Studies	Aula Global	
26/04	Models for Panel Data	R Case Studies	Aula Global	PS 5 (Coding)
04/05	Power and Other Topics	R Case Studies	ch. 16-17	
11/05	Causal Inference	R Case Studies	ch. 18-22	PS 6 (Theory)

Note: Schedule may be subject to change depending on our progress during the semester.

Additional required and/or recommended readings will be available on Aula Global

Note on Assigned Readings: Additional required and/or recommended readings will be available on Aula Global. Whether you complete readings before the associated lecture session or after is up to you. You should complete them in relatively close to the class session associated with a given topic in the readings. Most of the class sessions will consist of lectures, and mostly on theory, concepts, and simple examples. In general, readings will cover more material than the associated lecture. Despite the lecture format the nature and goals of this class require active involvement, discussion, and participation from students.

IV Grading and Work Load

Your final grade will be determined based on the following three components:

1. **Problem sets (**$40\% = 4 * 10\%$**):** The main focus of this course will be your assignments. I strongly encourage you to work in groups and discuss each question with your peers. However, each student must write up and submit their own original solution. Problem sets have to be submitted via Aula Global by the specified due date (before the start of our lecture). Three problem sets will focus on coding challenges, while the remaining three will cover theoretical and conceptual questions. *Of the 6 problem sets, I will take the average of the 4 highest grades (2 theory and 2 coding).*

2. **Research project (**30%**):** At the end of the semester, you are expected to submit a research project that utilizes the methods and techniques covered throughout the course. While you have to incorporate an original data analysis using R, you are free to choose any topic and/or data source you find interesting (and it may overlap with your other substantive coursework). Further details will be discussed in class.

3. **Final exam (**30%**):** The final exam will test you on all the material covered throughout the semester. It will focus on the theoretical questions related to the statistical models discussed in the lecture.

V Acknowledgements

I have adapted the ideas and language from the work of several educators for this syllabus and the course material. For example, I have borrowed liberally from other courses on social science research methods and statistics, as taught by Andrew Gelman, Gary King, Michael Peress, Thomas Gschwend, and others. I appreciate their contributions to the discipline and thank all educators who make their teaching material available to others. To pay it forward, I will share my own material with anyone who is interested.

References

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Chang, Winston. 2018. *R graphics cookbook: practical recipes for visualizing data*. O'Reilly Media.

Fox, John. 2015. *Applied regression analysis and generalized linear models*. 3 ed. Sage Publications.

Fox, John, and Sanford Weisberg. 2018. *An R companion to applied regression*. 3 ed. Sage Publications.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and other stories*. Cambridge University Press.

Grolemund, Garrett. 2020. "The Tidyverse Cookbook."

Healy, Kieran. 2018. *Data visualization: a practical introduction*. Princeton University Press.

King, Gary. 1998. *Unifying political methodology: The likelihood theory of statistical inference*. University of Michigan Press.

Long, J Scott. 1997. *Regression models for categorical and limited dependent variables.* Thousand Oaks: Sage Publications.

Wickham, Hadley, and Garrett Grolemund. 2016. *R for data science: import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.

Wilke, Claus O. 2019. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media.

Wooldridge, Jeffrey M. 2013. *Introductory econometrics: a modern approach*. Cengage Learning.