MES 624: Data Analysis for Environmental Policy

Renato Molina

Fall 2019

E-mail: renato.molina@rsmas.miami.edu
Office Hours: T 10:00-12:00
Office: CIMAS 213
Teaching Asisstant: Madeline Kaufman

Web: renatomolinah.com/teaching
Class Hours: M/W 12:00-1:15
Class Room: MSC 123
Section Hours: M 3:00-4:00

UM Bulletin Description

The objective of this class is to provide students with applied skills that will allow them to understand, discuss, and perform general quantitative analyses related to environmental policy. The course will cover required statistical background and methods, as well as an introduction to advanced topics and challenges in population and causal inference. Students will also be trained on how to use computing programming tools to manipulate data and perform analyses. The class is intended for graduate students seeking to learn modern methods and techniques for quantitative analysis of environmental problems and their application in policy making.

Course Description

The objective of this class is to provide students with quantitative skills to understand, discuss, and perform quantitative analyses with environmental data. Emphasis will be set on building research and policy intuition that they can then apply in a wide range of environmental problems.

The course will cover required statistical background, as well as an introduction to advanced topics to perform population and causal inference in environmental settings. At the end of the class, the students will be proficient in descriptive statistics, probability, and the fundamental concepts and applied methods in statistical inference and regression analysis. In addition and through weekly discussions, students will learn how to use computing programming tools to put these concepts into practice. Students are expected to become fluid in coding tasks, data handling, and the communication of their quantitative analyses.

The class is intended for graduate students seeking to familiarize with modern quantitative methods and techniques for analysis of environmental data. It is expected that students taking this class have a basic to medium command of statistics and algebra at the undergraduate level.

Prerequisites: None.

Course Objectives

1. Students will learn the fundamental concepts and applications associated with descriptive statistics, probability, and probability distributions.
2. Students will learn the fundamental concepts and applications associated with statistical inference. This set of concepts include hypotheses testing, test of goodness of fit, analysis of variance, and nonparametric testing.
3. Students will learn the fundamental concepts and applications behind regression analysis and causal inference. This set of concepts include ordinary least-squares, properties of estimators, fixed-effects, and challenges in regression analysis.
4. Students will learn how to use computer programming to apply these concepts in environmental problems.

Readings

Required

Ramsey, Fred and Daniel Schafer (2012). *The statistical sleuth: a course in methods of data analysis*. Cengage Learning.

Optional

Angrist, Joshua D and Jörn-Steffen Pischke (2014). *Mastering'metrics: The path from cause to effect*. Princeton University Press.

Angrist, Joshua and Jorn-Steffen Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton University Press.

Ashenfelter, Orley C, David Zimmerman, and Philip Levine (2002). *Statistics and Econometrics from A to Z*. John Wiley & Sons.

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

Advanced

Casella, George and Roger L Berger (2002). *Statistical inference*. Vol. 2. Duxbury Pacific Grove, CA.

Larsen, Richard J, Morris L Marx, and others (1986). *An introduction to mathematical statistics and its applications*. Vol. 2. Prentice-Hall Englewood Cliffs, NJ.

Course Policies

Grading Policy

- Assignments (**20%**)

- Two midterms (**25%** each)
- A final examination (**30%**)

Four assignments will be posted throughout the semester. All of the assignments are in the form of a problem set. The deliverable is a report documenting all the work associated with their answers in groups of two to three students. The students are expected to produce high quality, polished, professional work. Late work will not be graded, unless an exception has been granted prior to the due date.

Assignments are key to mastering the material and preparing for the examinations. During discussion, the teaching assistant will work with students to cover the necessary skills to solve the assignments using the concepts covered in class as well as the available computing tools.

All midterm and final examinations will be in the form of an in-person written test. In the case of professionally documented special examination needs, proper individual accommodations will be granted.

Attendance Policy

Students are expected to abide by the University of Miami's attendance policy. Other absence reasons should be discussed with me in advance.

E-mail Policy

I am usually quick to respond to e-mails. Sometimes, however, I may get busy or on the road and fail to write back in time. If you do not receive a response after one or two days, please feel free to reach out again.

Honor Policy

All students are expected to abide by the University of Miami's honor code. Anyone caught cheating on exams, improperly referencing published written or electronic material, or submitting work that is not their own will fail the course.

Technology During Lectures Policy

The use of technology for any other purpose that is not taking notes or working on the class content is highly discouraged. Repeated violations to this policy will result in final grade penalizations.

Lecture Schedule

Students must read the assigned chapters or papers before lecture. Important: Class content and readings are subject to change, contingent on mitigating circumstances and the progress of the class. Any changes will be announced via Blackboard.

Week 01, 08/19 - 08/23: Introduction and Statistical Inference

Ramsey & Schafer Ch. 1 • Ashenfelter, Zimmerman & Levine Ch.1

- Statistical inference and study design
  - Causal inference
  - Inference to populations
  - Statistical inference and chance mechanisms
- Measuring uncertainty in randomized experiments
- Measuring uncertainty in observational studies

Week 02, 08/26 - 08/30: Inference Using t-Distributions

Ramsey & Schafer Ch. 2 • Ashenfelter, Zimmerman & Levine Ch.8

- One sample t-tools and paired t-Test.
  - Distribution of sample averages
  - t-ratio on sample averages
- t-Ratio for two-sample inference
  - Distribution of the difference between two independent sample averages
  - Confidence interval for the difference between two independent sample averages
  - Testing hypotheses about the differences in means
  - P-value
- Inference in a two-treatment randomized experiment

Week 03, 09/02 - 09/06: Asssumptions in t-Models

Ramsey & Schafer Ch. 3 • Ashenfelter, Zimmerman & Levine Ch.7

- Robustness of the two-sample t-tools
  - Departures from normality
  - Differing standard deviations
  - Departure from independence
- Resistance of the two-sample t-tools
  - Outliers and resistance
- Strategies for two-sample problems
- Transformation of the data

Week 04, 09/09 - 09/13: Alternatives to t-Tools

Ramsey & Schafer Ch. 4 • Ashenfelter, Zimmerman & Levine Ch.6

- Tests for two-independent samples
  - Rank sum, permutation, and Welch t-test
- Tests for paired data
  - Sign, and Wilcoxon-signed-rank test

Week 05, 09/16 - 09/20: Multiple Samples

Ramsey & Schafer Ch. 5 • Ashenfelter, Zimmerman & Levine Ch.5

- Pooled sample comparison
- One-way analysis of variance F-test
  - Extra-sum-of-squares principle
  - ANOVA table for one-way-classification
- Robustness and model checking

Midterm 1

Week 06, 09/23 - 09/27: Linear Combinations and Multiple Comparisons

Ramsey & Schafer Ch. 6 • Ashenfelter, Zimmerman & Levine Ch.5

- Inference in linear combinations
- Simultaneous inferences
  - Tukey-Kramer procedure
  - Scheffeé's procedure

Week 07, 09/30 - 10/04: Simple Linear Regression

Ramsey & Schafer Ch. 7 • Ashenfelter, Zimmerman & Levine Ch.9

- The simple linear regression
- Least Squares regression estimation
  - Fitted values and residuals
  - Least squares estimators
  - Sample distribution of the least-squares estimator
- Tools for inference
  - Tests and confidence intervals
  - Prediction and calibration

Week 08, 10/07 - 10/11: Extensions in Simple Linear Regression

Ramsey & Schafer Ch. 8 • Ashenfelter, Zimmerman & Levine Ch.10

- Robustness of least squares inferences
- Interpretation of transformations
- Duality with ANOVA
  - Population means
  - Lack-of-fit F-test
  - Composite analysis of variance

Week 09, 10/14 - 10/18: Multiple Regression

Ramsey & Schafer Ch. 9 • Ashenfelter, Zimmerman & Levine Ch.11

- Construction and interpretation of regression coefficients
- Constructed variables
  - Squared terms
  - Indicator variables
  - Categorical variables
  - Interacted variables
- Strategy for data analysis

Week 10, 10/21 - 10/25: Extensions in Multiple Regression

Ramsey & Schafer Ch. 10 Salvatore & Reagle Ch.8 • Ashenfelter, Zimmerman & Levine Ch.12

- Inference for multiple coefficients
- Extra-sums-of-squares F-tests
  - Residuals in hierarchical models
  - F-test for joint significance
  - ANOVA

Midterm 2

Week 11, 10/28 - 11/01: Two-Way ANOVA

Ramsey & Schafer Ch. 13

- Additive models
- Saturated, non-additive models
- F-test for additivity
- Randomized blocks and paired-t analysis

Week 12, 11/04 - 11/08: Logistic Regression

Ramsey & Schafer Ch. 20-21 • Ashenfelter, Zimmerman & Levine Ch.16

- Inference in logistic regression
  - Maximum likelihood
  - Test and confidence intervals
- Probit regression
- Binomial responses
  - Wald test
  - Extra-binomial variation

Week 13, 11/11 - 11/15: Specification Error, Multicollinearity, and Measurement error

Ramsey & Schafer Ch. 3,11 • Ashenfelter, Zimmerman & Levine Ch.13

- Ommited variable bias

- Nonliearities
- Multicollinearity
- Measurement error
  - Probability limits
  - Errors-in-variables
  - Instrumental variables

Week 14, 11/18 - 11/22: Heteroskedasticity, Serial Correlation, and Simultaneous Equations

Ramsey & Schafer Ch. 15 • Ashenfelter, Zimmerman & Levine Ch.14-15

- Heteroskedasticity
  - Consequences for regression analyses
  - Detection
  - Weighted-lest-squares
  - White correction
- Serial correlation
  - Consequences for regression analyses
  - Detection
  - Quasi-differencing
  - Durbin-Watson test
- Simultaneous equations
  - Simultaneous equations bias
  - Identification
  - Order condition
  - Indirect and two-stage least squares

Week 15, 11/25 - 11/29: Thanksgiving Recess

Week 16, 12/02 - 12/06: Review and Final

Final Examination

Discussion Schedule

Week 01, 08/19 - 08/23: R-Studio

Week 02, 08/26 - 08/30: Functions and loops

Week 03, 09/02 - 09/06: Simulation

Week 04, 09/09 - 09/13: Optimization

Week 05, 09/16 - 09/20: Data Wrangling

Week 06, 09/23 - 09/27: Differences in Population Mean

Week 07, 09/30 - 10/04: Regression

Week 08, 10/07 - 10/11: Prediction

Week 09, 10/14 - 10/18: Multiple Regression Analysis

Week 10, 10/21 - 10/25: Inference from observational studies

Week 11, 10/28 - 11/01: Inference from randomized experiments

Week 12, 11/04 - 11/08: Survival studies

Week 13, 11/11 - 11/15: Instrumental variables

Week 14, 11/18 - 11/22: Two-Stage Least Squares

Week 15, 11/25 - 11/29: Thanksgiving Recess

Week 16, 12/02 - 12/06: Review