Course objectives: The goal of this course will be to cover methods for making causal inferences in data. Such approaches include discussion on analyzing experimental data and observational studies. Experimental approaches are introduced first, and then it is shown how much analysis of observational data attempts to “mimic” traditional experimental design.

Computing: SAS and R. Students will be provided with programs and code to conduct all analyses that are discussed. Additionally, each procedure will be demonstrated when presented.

Course Schedule

Day 1: Reviewing Foundational Material
- Review Matrix Algebra
  o Notational Convention (summation, products, parameters, statistics, etc.)
  o Algebraic operations and rules
  o Matrix operations (transpose, trace, etc.)
  o Eigenvalues and eigenvectors
  o Determinants/Generalized Variance
  o Geometric Interpretation of Matrix Operations
- Review Multivariate Normality
  o Tests of Multivariate Normality (Mardia’s test, etc.)
  o Visualization Methods
    • Chi-square Plots
    • Density Plots
- Introduce the R software

Day 2: Experimental Design
Under experimental design, several different traditional experimental methods are discussed, including:
- Single Factor, One Way Analysis of Variance
  o Analyzed with the Multivariate Normal Assumption
  o Analyzed from a Randomization Perspective (e.g., Lady Tasting Tea)
- Two-Way Analysis of Variance
- One-Way Multivariate Analysis of Variance
- Demonstrate ANOVA models in SAS and R

Day 3: Experimental Design and the General Linear Model
- Two-way Multivariate Analysis of Variance
  o Parallel Analysis
  o Coincidence Analysis
- Multivariate Analysis of Covariance
  o Review Assumptions of MANCOVA
- Review of General Linear Models
- Multivariate Multiple Regression (showing how Experimental Design is a Special Case)
  o Assumption Checking
  o Residual Analysis
  o Piecewise Regression
- Demonstrate MANOVA, MANCOVA, GLM in R and SAS
Day 4: Inference in Observational Studies
  - Regression Discontinuity Design
    o Standard RDD Design
    o Fuzzy RDD Design
    o Inference in RDD
  - Propensity Scores
    o Estimation via Logistic Regression
    o Matching Methods
      • Nearest Neighbor
      • Optimal Matching
    o Propensity Scores as Weights
    o Propensity Scores in Multilevel Models
      • Includes short review of Multilevel Modeling
  - Demonstrate RDD and Propensity Scores in R

Day 5: Inference in Observational Studies Continued
  - Using Instrumental Variables (LATE)
    o Show relationship with SEM and Latent Variable Modeling
  - Differences-in-Differences Regression Designs
    o Show Relationship to Interactions in MANCOVA model
  - Demonstrate LATE and D-i-D in R