

Guidelines

Midterm & Practicum 1

Prof. Jack Reilly

Fall 2026

There are two exams that take place around the mid-point of the semester: (1) an in class “midterm” exam and (2) a take-home “practicum” exam.

- On the **in-class “midterm” exam**, you are permitted the full class time (although the exam will likely not take the whole time). You may bring your own simple calculator (not graphing). If you do not have one, the department will provide one. You need to bring a pen and/or pencil. You may also bring a “cheat sheet” — a single page 8x11 sheet of paper with whatever you want written (or typed) on the page.
- On the **take-home “practicum” exam**, you are permitted as much time as you like between when the exam is handed out and when it is due. The take-home IS **open-book**, **open-note**, and **open-internet**, but it IS NOT **open-human** or **open-ai**. In short: you can use any resource you want, so long as that resource does not involve asking another intelligence, whether human or artificial, a question.
- You can ask the professor questions.
- On both tests, all material covered in the class, in reading, or in homework is fair game, although the take-home will lean towards using computation and the in-class portion will lean towards conceptual questions.

The [reference tables](#) page shows the statistical tables you will have available to you on the exam.

Review Terms and Concepts by Unit

1. Basics

- Data
- Cases
- Value
- Unit of Analysis

- Estimator
- Two main goals of data analysis
- Hypotheses
- Dependent Variable
- Independent Variable
- Population
- Sample
- Parameters
- Statistics
- Descriptive Statistics
- Inferential Statistics

2. Descriptive Statistics

- Kinds of variables
- Measurement metrics
- “Exploratory data analysis”
- Kinds of univariate graphics (histogram, box plots, etc)
- Tabulations and crosstabulations
- Shape, center, spread, skew
- Mistakes with graphics: axes, scales, etc
- Measures of central tendency
- Measures of dispersion
- Outliers
- Linear transformations
- Density curves
- Normal distribution
- Z-scores
- Cumulative probabilities

3. Statistical Relationships

- Kinds of bivariate graphics (scatterplots, linear fits, smooth lines, etc)
- Scatterplot diagnoses: form, direction, strength, outliers
- Transformed data in scatterplots
- Correlation coefficients
- Sample vs. population calculation differences
- Regression
 - intercept
 - slope
 - model component
 - stochastic component

- sources of error in modeling
- method of least squares
- model sum of squares
- residual/error sum of squares
- total sum of squares
- R^2
- $\hat{Y}$
- Y_i
- $\bar{Y}$
- Calculating the slope (b_1) from a correlation coefficient
- Regression coefficients vs correlation coefficients
- Residual plots
- Effects of outliers on regression lines
- Predicted values/expected values of Y given X
- Two-way tables for qualitative variables

4. Research Design

- Correlation vs. causation
- Confounding variables
- Conditions for causation
 - credible causal link (theory)
 - temporal precedence (could Y cause X?)
 - covariation (do X and Y move together/covary?)
 - no plausible alternative explanations (no confounding variables)
- Simpson's paradox
- Internal validity
- Threats to internal validity
 - History
 - Maturation/Learning
 - Testing
 - Instrumentation
 - Regression to the mean
 - Selection bias
 - Mortality
 - Social Desirability
- External validity
- Threats to external validity
 - context
 - sampling procedures
- Anecdotal data
- Survivor's bias
- Control through randomization
- Experiments

- treatment
- control
- observations
- random assignment
- Surveys
 - Population/sample/sampling frame
 - Simple random sample
 - Stratified random sample
 - Multi-stage sampling
 - Response rate
 - Response bias
 - Question framing and ordering effects
 - Bias vs. variability (validity vs reliability)

5. Human Subjects & Ethics

Not required.

- Infamous studies: Nazi medical experiments, Tuskegee syphilis study, Milgram obedience experiment, Stanford prison “experiment”
- Belmont report, National Research Act of 1974
- IRB (Institutional Review Board)
- Core ethical principles
 - Informed consent
 - Anonymity
 - Long-term harm

6. Probability

- Random phenomenon
- Independence
- Probability model
- Sample space
- Event
- Long run
- Rules of probability
 - Range
 - Sample space
 - Addition
 - Complementarity
 - Multiplication
- Disjoint events
- Random variable

- Discrete
 - Continuous
- Probability distribution
- Expected value (mean) of a random variable
- Law of large numbers
- Standard deviation/variance of a random variable
- Rules for means of random variables
 - Addition
 - Subtraction
 - Linear transformations
- Conditional probability
- Probability trees

7. Stata

- Be able to:
 - Load data
 - Use a .do file
 - Place comments in .do files
 - Calculate summary statistics
 - Graph scatterplots and histograms
 - Recode data
 - Tabulate and cross-tabulate data
 - Run regressions
 - Predict values from regression results

- Selected Stata commands

```
cd, use, clear, tab, sum, tabstat, gen, sort, browse,
tsway, scatter, histogram, graph export, corr, reg,
disp, bys, predict, save, keep, drop, if, recode,
collapse, preserve, restore
```