STAT 113, FALL 2017: SPECIFIC LEARNING OBJECTIVES

Below is a list of the Specific Learning Objectives (SLOs) that the course grade is based on. Each graded item (e.g., homework, quiz or exam question) is associated with one of more of these objectives. If the response to a graded item shows full or partial mastery on an SLO, it will be assigned a mark of "M" for "Mastery", "M-" for near-Mastery, or "P" for progress. A grade of N means the response was not sufficient to demonstrate understanding of the concept in question. For averaging purposes, Ms count as 4.0, M^- as 3.0, P as 2.0 and N as 0.0.

Concepts SLOs (50% of final grade)

Concepts SLOs concern understanding of foundational statistical ideas. The grade for each SLO is the average of the top up to two items pooled across quizzes, exams and the individual project, and the top item from homework if it is higher than the quiz/exam average. For example, for SLO B3 a student has an M and an M- on exams/quizzes, and an M- on homework. The exam/quiz average is 3.5 and the combined average is $(4 + 3 + 3)/3 = 3.33$. Since the exam/quiz average is higher, the grade is 3.5. If the M were on homework instead, the combined grade would be used, since this would be higher than the 3.0 quiz/exam average.

Reassessment Policy. You may request extra assessments ("reassessments") for a given Concepts SLO if you were not satisfied with your original mark on a quiz/exam/individual project, provided you have turned in all assigned homework related to that SLO (including ungraded problems). It does not have to be correct, but you need to have made a "good faith" attempt. SLOs missed initially by a large portion of the class may be revisited on later homeworks or quizzes.

Reassessments are requested by making an appointment with me to come to office hours and take a "reassessment quiz". You may be reassessed on a maximum of 2 Concepts SLOs per week; you cannot make up all of your SLOs in the last weeks of the semester!

Date: Last Revised August 18, 2017.

A. Structure of data

1. Recognize cases/observational units
2. Distinguish categorical from quantitative variables
3. Identify explanatory and response variables in a study design

B. Study Design

1. Identify and distinguish samples and populations
2. Recognize potential sources of bias in sampling methods
3. Distinguish observational and experimental designs, and understand when there is evidence for a causal relationship between two variables
4. Understand what is needed for a variable to be considered a "confounding variable".
5. Understand how controlling for a confounding variable can change the direction and/or strength of an association.

C. Describing and visualizing data

1. Understand properties of common measures of center (mean, median), and when to use each one.
2. Choose and calculate appropriate proportions from contingency tables to answer a question
3. Understand properties of common measures of variability (range, IQR, variance, standard deviation) and identify when it makes sense to use each.
4. Compare and describe distributions using common plots (histograms, box plots, density plots, etc.)
5. Describe and interpret relationships between quantitative variables using scatterplots.

D. Using and interpreting predictive models

1. Use simple linear regression models to make predictions, estimate their intercept and slope by eye from plots, and interpret the coefficients in terms of real-world quantities.
2. Use residual plots to evaluate whether a linear model is appropriate to summarize a relationship.

E. General principles of statistical inference

1. Identify and distinguish statistics and parameters.
2. Calculate and interpret standard measures of effect size.
3. Understand how a sampling distribution is created, and recognize what the cases represent and what quantity the distribution is over, and where they are (usually) centered.
4. Understand what standard error represents and how it is found using a sampling distribution.
5. Identify the aspects of a sample or population that affect the precision of an estimate.
6. Understand and convert between standardized and unstandardized statistics.
7. Identify when the conditions for analytic approximations are met.
8. Distinguish strength of evidence from size of effect.
9. Identify when it is and isn't appropriate to use a hypothesis test and a confidence interval.
10. Choose appropriate estimation/testing procedures to answer particular questions.

F. Using and interpreting confidence intervals

1. Interpret confidence intervals and the meaning of the confidence level
2. Understand how bootstrap distributions are created, recognize what the cases represent, what quantity the distribution is over, and where they are centered.

G. Principles of Hypothesis Testing

1. Identify appropriate null and alternative hypotheses corresponding to a research question.
2. Understand how randomization distributions are created, recognize what the cases represent, what quantity the distribution is over, and where they are (usually) centered.
3. Interpret P values as proportions and/or probabilities.
4. Interpret P values as measures of evidence.

5. Understand the kinds of statistical errors that can occur and what they mean in context.
6. Understand how aspects of the sample and population affect likelihood of false discoveries (Type I Errors) and missed discoveries (Type II Errors).

Computing (10%)

Although the concepts are the most important part of learning statistics, computing is a major component of virtually every facet of modern statistical practice from data manipulation to visualization to calculation to simulation, and as such it is important to be proficient in representing research questions and statistical methods in explicit computational terms. The Computing SLOs below are assessed on labs and the individual mini-project. The top up-to-three graded items for each SLO count toward the grade.

H. Statistical technology

1. Use software to create, format, and manipulate datasets to prepare them for the relevant analyses
2. Create informative visualizations using software
3. Implement randomization and simulation methods in software
4. Fluently use standard commands and interpret output from standard software
5. Produce well documented and reproducible lab reports

Methods SLOs (20%)

Two of the main goals of the class are (a) estimating quantities and characteristics of populations and (b) testing hypotheses about those quantities or characteristics using samples. You will be asked to accomplish both goals (estimation and testing) in the form of somewhat open-ended "guided investigations" (in homework, exam problems, and the mini-project) in the service of a real research question. In your investigations you should always begin with appropriate visualization(s) of the data, followed by numerical description, estimation, and testing techniques that best shed light on the research question.

I. Investigations using Simulation
   1. A single proportion
   2. A difference of proportions
   3. A single mean
   4. A difference of two means

J. Investigations using Analytic Methods
   1. A single proportion
   2. A difference of proportions
   3. A single mean
   4. A difference of two means
   5. A mean of differences from paired samples
   6. A simple linear regression model
   7. Fit of a set of proportions to a specified model (goodness of fit test)
   8. Association between two categorical variables
   9. Differences among more than two means (analysis of variance)

Project SLOs (20%)

The group project is graded using the following 25 criteria, each graded on the $M/M^-/P/N$ scale. The four graded components are (a) the proposal, (b) the initial analysis of the pilot data, (c) the final presentation, and (d) the final writeup. Criteria associated with more than one component (as listed in parentheses below) can be improved in revision.

K. Data-Collection

1. (Pilot, Final) Identify appropriate statistical hypotheses based on the research question.
2. (Proposal, Pilot, Final) Design an appropriate sampling procedure to provide evidence pertaining to the research question.
3. (Proposal, Pilot, Final) Select appropriate variables to measure.
4. (Pilot, Final) Identify limitations of the data-collection procedure and/or the variables measured, indicate how these constrain interpretation of the results, and identify ways the study could be conducted differently to imporove interpretability.

L. Descriptive Statistics

1. (Pilot, Final) Select appropriate visualizations to illustrate key properties of the dataset with respect to the research question.
2. (Pilot, Final) Describe what visualizations suggest about questions of interest.
3. (Pilot, Final) Select appropriate descriptive statistics to illustrate key properties of a dataset with respect to a question.
4. (Pilot, Final) Interpret descriptive statistics in the context of the question.

M. Inferential Statistics

1. (Pilot, Final) Select appropriate inferential methods based on the research question.
2. (Pilot) Carry out technically correct inferential procedures.
3. (Final) Carry out technically correct inferential procedures.
4. (Final) Assess any necessary conditions for inference.
5. (Pilot) Interpret inferential results in context.

6. (Final) Interpret inferential results in context.

N. Computing

1. (Pilot, Final) Data set is well prepared and formatted for analysis
2. (Pilot) Supporting code is readable, complete and consistent with reported results
3. (Final) Supporting code is readable, complete and consistent with reported results
4. (Pilot) Simulation methods are used and carried out in R code.
5. (Final) Analytic methods are used and carried out in R code.
6. (Pilot, Final) Paper is in the form of a fully self-contained and reproducible RMarkdown report

O. Communication

1. (Proposal, Pilot, Final) Clearly explain why the research question is of interest, and place the current study in context of previous research.
2. (Pilot) Clearly explain and justify the methods used.
3. (Final) Clearly explain and justify the methods used.
4. (Presentation) Oral presentation is clear, concise, and contains informative visual aids.
5. (Pilot, Final) Motivation, results, and interpretation are presented clearly and concisely in written natural langauge.