Statistics 411/511
Homework 1
Due Tuesday, October 5 by midnight

- **Instructions:** Please see the end of the syllabus for guidelines. Upload your homework to Gradescope via Canvas (access specific homework assignments from the "Assignments" link at the left of the Canvas course page). Your file must be a pdf document, and please be sure to assign pages. (See this Gradescope help video).
- Do the computational part of the homework shortly after completing the week's lab activity.
- The problems are assigned from the **third edition** of the textbook. If you have another edition, consult the copy on one-hour reserve at the library website for the homework problems.
- **Academic Integrity** You are encouraged to *discuss* the homework with other students, but what you turn in must be your own work in your own words. **DO NOT** copy someone else's homework. You may share ideas and R code, but do not share R output or written language. The syllabus contains details and links to OSU's Student Conduct Code and procedure for reporting suspected academic misconduct.

1. A study compared new COVID infection rates between communities that had instituted lockdown policies and communities that had not instituted lockdown policies. The new COVID infection rates were lower in the communities that had gone into lockdown. Can the researchers conclude that lockdowns cause lower new COVID infection rates? Explain briefly.

2. Last fall, OSU conducted TRACE OSU, a study of COVID-19 prevalence at the OSU campuses in Corvallis and Bend, as well as at the Hatfield Marine Science Center in Newport. Below is a description of how subjects were selected for the study:

   > Current faculty, staff and students were invited to participate in this testing by enrolling in the study via a webform. Each week during the academic year, up to 1,000 names of faculty, staff and students, who have enrolled to participate in TRACE OSU and who reside in the Corvallis, Bend or Newport areas were drawn randomly from the pool of enrollees and invited by e-mail to be tested.

   There is only one group, so there can be no random allocation to groups. The only issue relevant to scope of inference is the selection of study units. Write a brief statement (one or two sentences) describing the scope of inference for the TRACE OSU COVID-19 study.

3. The data described in exercise 12 on page 23-24 are available in the Sleuth3 R package. It's called `ex0112`. Consider these research questions:

   i Does the fish oil diet affect diastolic blood pressure differently than the regular oil diet?
   ii What is the difference in diastolic blood pressure reduction between the two diets?

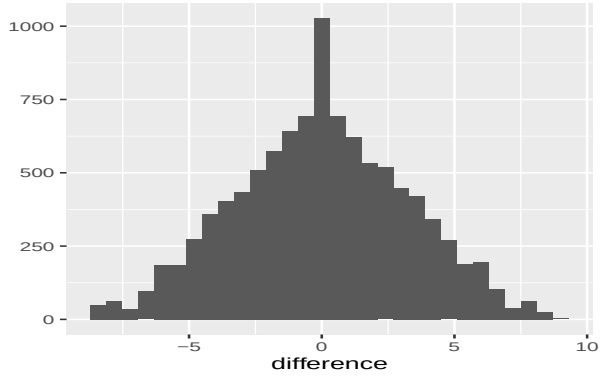
   Using the data in `ex0211`, do the following.

(Problem 3 continued on next page)

(a) Produce side-by-side boxplots for the two groups. Include your plot and your R code.

(b) Obtain summary statistics for each diet group. Include your R code and output.

(c) I did a permutation test by calculating the difference in sample means for 10,000 random shuffles of the data into two groups of size 7 (same sample sizes as the observed two diet groups). Below is a histogram of the difference in sample means. Describe where the observed value of the difference falls on this histogram, and comment on the plausibility of the null hypothesis of no difference in diastolic blood pressure reduction between the two diet groups.



(d) Perform a two-sided t-test using R's `t.test()` function. Turn in your R code, but not the output. Instead, write a brief "Statistical Conclusion" similar to those given for the two case studies in Chapter 1, answering the two research questions above. See the end of the syllabus on Canvas for detailed instructions on writing statistical conclusions for hypothesis tests and confidence intervals.