Statistics 411/511
Homework 4
Due Tuesday, October 26 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. **There will be a one-point deduction if you don't 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. Revisit the blood pressure data from question 3 of Homework 1, `ex0112`. The research questions are similar to those in Homework 1: "Does the fish oil diet reduce diastolic blood pressure more than the regular oil diet, and if so, by how much?"
   (a) Make side-by-side boxplots of the reduction in diastolic blood pressure for the two diet groups. Include your plot but not your R code.
   (b) Does your plot in part (a) suggest that the equal variance assumption for the two-sample t-test is violated? Why or why not?
   (c) Does your plot in part (a) tell you anything about the independence assumption for the two-sample t-test? (yes/no)
   (d) State the null and alternative hypotheses to answer the research question with a two-sample t-test. Your hypotheses should be in terms of population parameters. Define any notation you use for the population parameters.
   (e) Conduct a Welch's t-test of your hypotheses in part (d). Include your R code but not output.
   (f) Write a statistical conclusion to report the results of your hypothesis test in part (e).
   (g) Obtain a two-sided confidence interval for the difference in population means. Do not submit R code or output. Instead, write a statistical conclusion reporting this interval.

2. This problem deals with the data of exercise 3.18 on page 78 of the textbook. The data frame is not in the `Sleuth3` package. You can use the following R code to create ex0318:

```
> ex0318 <- data.frame(
+   Metabol=c(20.1, 22.9, 18.8, 20.9, 20.9, 22.7, 21.4,
+             20, 38.5, 25.8, 22, 23, 37.6, 30, 24.5),
+   Group=c(rep("Nontrauma", times=8), rep("Trauma", times=7)))
```

(Problem 2 continued on next page)

(Note that the three +'s are R continuation prompts. Don't type them.)

(a) The research question is whether metabolic expenditure is higher in trauma patients than in non-trauma patients.

State null and alternative hypotheses in terms of the shift parameter discussed in item 2(g) of Lab 4 and on page 3 of Outline 4.

(b) Conduct a Wilcoxon rank-sum test of your hypotheses in part (a). Submit your R code but not output.

(c) Write a statistical conclusion for the hypothesis test in part (b).

3. Revisit the cat data of Homework 2. In problem 1 of Homework 2, you computed a paired t-statistic to test the null hypothesis that the mean difference is 0. Use the sign test to test the null hypothesis that the median difference is 0 vs. the alternative that the median water consumption is different for still water vs. flowing water. Include your R code and output.