Problem Set 9
STAT-S 520

Due on March 27th, 2023

Instructions:

- Submit your answers in Canvas.
- Your answers can be typed and/or handwritten as long as your final submission is a single PDF file with answers in proper order.
- Include your R code, graphs, and output. The latter only when is relevant.

  - Check that only the relevant output is included in your submission. Pages and pages of output that is not relevant can be penalized.

- You are allowed to collaborate with your classmates as long as you write your own solutions.

Questions:

For all the questions, use a significance level $\alpha = 0.025$.

1. Redo ISI Section 9.6 Exercise 5, but do the following:

    i. Solve (c) and (f) using the simulation-based approach.
    ii. Assume instead that a random sample of $n = 100$ chickens reveals 4 diseased chickens. Use the traditional theory-based approach.
    iii. Redo (ii) using the simulation-based approach. How does that compare to the results obtained in (ii)

2. ISI Section 9.6 Exercise 7 using the theory-based approach. In addition, can you perform a simulation-based approach for this problem? If YES, do it. If NO, explain why not.

3. Use the variable `arr_delay` (arrival delay in minutes) from the data frame `flights` from the package `nycflights13` (you need to install this package in R first)

    a. Treat the variable `arr_delay` as the population of interest. Create an object that contains the information of `arr_delay` without missing values (use `na.omit( )` to remove missing values)
    b. Study the distribution of your sample. Does the data seem to be drawn from a normal distribution?
    c. Use `set.seed(520)` right before getting a sample of 100 arrival delays. Perform a test to determine whether the average arrival delay for NY flights is less than 10 minutes? State your hypotheses, test statistic, p-value, and conclusion.
    d. Using the same sample as in part c, perform a test to determine whether more than 50% of flights have no arrival delays. State your hypotheses, test statistic, p-value, and conclusion.

4. Do the following:

    a. Redo question 3c using the simulation-based approach.
    b. Redo question 3d using the simulation-based approach.

Reading assignments

- ISI Chapter 9, Section 9.5
- ISI Chapter 10, Section 10.1