Test 2

One- and two-sample inference for means [L3, L4, L5]

STAT218

2024-05-10

## Instructions

You have 48 hours from the release of this assignment to complete and submit your work. You may refer to all class materials, notes, and textbooks, but must complete this assignment on your own. By submitting your work, you are affirming that your work is your own and you have not consulted with anyone else in preparing your answers or generated your answers or analyses using AI. Failure to adhere to this expectation will be considered an act of academic dishonesty and result in loss of credit.

You will find a project with a mostly empty script in the class Posit cloud workspace; use this to complete your analyses where required. Note that not all parts require you to perform any calculations; some questions are purely qualitative. Use the prompts as your guide, not the script.

Once you have completed your analyses for the portions requiring use of statistical software, submit your work by filling out the [test 2 form] (also posted on the course website). The form will automatically save your work, so you can return to it over the course of the 48-hour test window.

The form will stop accepting responses at the deadline, so **make sure you submit by 5pm on Friday 5/10**. Lastly, keep in mind that you will be given the opportunity to revise problems that you miss the first time around to earn back credit.

## Problems

### Problem 1: gifted children

1. [L3, L4, L5] The gifted dataset contains data on 36 children identified as gifted in a large city. Assume for the purpose of the problem that the data are from a random sample of gifted children in the metropolitan region where the data were collected.
   1. [L3] Is there a relationship between the mother’s IQ and the child’s test score? Construct a scatterplot and compute and interpret the correlation.
   2. [L3] Repeat but with the father’s IQ.
   3. [L5] Given your results in (a)-(b), formulate a hypothesis about which parent’s IQ is higher. Explain your reasoning.
   4. [L5] Construct a histogram of the pairwise differences between the mother’s IQ and father’s IQ for each child in the dataset and check the assumptions for inference using the model.
   5. [L4, L5] Test the hypothesis you proposed in (c) at the 1% level and provide a corresponding interval estimate. Report your results in the narrative style introduced in class.
   6. [L4, L5] It’s thought that the mean age by which infants can count to 10 is around two years old. Test the hypothesis at the 5% level that gifted children do this sooner and provide a corresponding interval estimate. Interpret the test results and interval estimate in context following the narrative style introduced in class.

### Problem 2: lizard running speeds

1. [L3, L4, L5] The dataset lizards contains running speeds measured in a laboratory race track for two species of lizards, Western Fence (WF) and Sagebrush (S) lizards. Speeds are recorded in meters per second.
   1. [L3] Construct side-by-side boxplots of top speed by species. Comment on whether the assumptions for inference using the model seem appropriate.
   2. [L4] Compute point estimates and standard errors for the mean top speed for each species.
   3. [L4] Compute and interpret 99.5% confidence intervals for the mean top speed for each species.
   4. [L4, L5] Test for a difference in mean top speed between species at the 1% significance level and provide an interval estimate at the appropriate confidence level. Interpret the test and estimate following the narrative style introduced in class.

### Problem 3: self- and cross-fertilization and plant vigor

1. [L3, L4, L5] Does self-fertilization produce less vigorous plants than cross-fertilization? The dataset plants contains measurements of plant heights in inches for 15 pairs of plants of the same age; one plant in each pair was grown from a seed from a cross-fertilized flower, and the other was grown from a seed from a self-fertilized flower.
   1. [L3] Visualize the distribution of differences in plant heights between the cross-fertilized and self-fertilized individuals. Does the plot alone suggest an answer to the question of interest?
   2. [L4, L5] Test, at the 2% level, whether mean height of plants grown from cross-fertilized seeds exceeds that of plants grown from self-fertilized seeds and provide a confidence bound for the difference at the level corresponding to your test. Report the results of your analysis in context following the narrative style introduced in class.

### Problem 4: creativity and motivation

1. [L1, L2, L3, L4, L5] The creativity dataset contains data from an experiment on the effect of intrinsic vs. extrinsic motivation on creativity. A random sample of 47 creative writing students at an unnamed university were randomly assigned to one of two groups, extrinsic and intrinsic; each subject was instructed to write two short poems, but those in the extrinsic motivation group were primed on the task in a way that oriented them to external motivations for writing, and those in the intrinsic group were primed on the task in a way that oriented them to internal motivations for writing. Poems were scored by judges for creativity on a 40-point scale, and each subject received an average score.
   1. [L1] What is the study population? Based on the study description, is the sample representative, and if so, why?
   2. [L2] What type of study is this? Based on the study description, can the data support causal inferences about motivation and creativity, and if so, why?
   3. [L3] Construct an appropriate graphical summary comparing the distributions of average scores by treatment group.
   4. [L4] Provide point estimates and standard errors for the mean creativity score in each group.
   5. [L5] Test the hypothesis that motivational framing has no effect on creativity at the 1% significance level. Compute an interval estimate for the difference in mean scores at the level corresponding to your test. Use your results from (c)-(d) to check assumptions. Write a short narrative summary of your results following the style introduced in class. (Don’t forget to include a point estimate and standard error.)