**Terms, topics, or concepts you should be familiar with:**

p-value

reproducibility crisis

statistic

parameter

continuous variable

discrete variable

nominal variable

ordinal variable

experimental vs observational studies

blinding

pseudoreplication

biological vs technical replicates

confounding variables

common faults in plots

sample

population

transformation

parametric

non-parametric

**R skills you should have**

Create matrices and vectors

Subset a vector, dataframe, or matrix to select only specific elements

Read a csv file to import data

Make a publication quality plot of 1, 2, or 3 variables that have a mix of continuous and discrete values

Perform the statistical tests mentioned below

**Tests you should be able to run**

Binomial

Chi-square

T-test (single sample, two sample, paired)

Permutation

Correlation test

**Example Problems**

You administer a vaccine candidate for covid-19 and a placebo each to 1000 individuals. You find that after 2 months there have been covid-19 cases in 32 placebo group and 6 vaccine group participants.

Does this vaccine work? \_\_\_\_\_\_\_\_\_\_\_

What test did you use? \_\_\_\_\_\_\_\_\_\_\_

What p-value was associated with this test? \_\_\_\_\_\_\_\_\_\_\_

You measure height of students at the MSC and the gym. Are the heights you measured significantly different

MSC: 126, 164, 148, 120, 178, 183

Gym: 151, 109, 151, 174, 118, 136

What test did you use for this question? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What p-value was associated with this test? \_\_\_\_\_\_\_\_\_\_\_\_\_

What do you infer from your test? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

You grow plants with two different potting soils and measure height at 21 days. Use a non-parameteric test evaluate the results of your study.

Soil1: 23, 12, 45, 23, 21, 45, 21

Soil2: 35, 45, 21, 34, 67, 23, 16

You measure reproductive success of fish in your study you record sex and number of offspring that survive to adulthood for each fish in the study. Use a permutation test to determine whether males and females have equal variance in reproductive success.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sex | f | m | f | f | m | f | f | f | m | m | m | m | m | f | m | m | f | m | f | f |
| Total reproduction | 10 | 16 | 9 | 8 | 0 | 9 | 9 | 9 | 8 | 3 | 17 | 14 | 17 | 8 | 9 | 18 | 8 | 7 | 7 | 8 |

Determine which species don’t have a 50:50 sex ratio

Species 1 6 males / 1 female

Species 2 345 males / 302 females

Species 3 127 males / 94 females

Use the iris data in R - data(iris) - to make a plot showing what you think is the most striking differences in the species