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

p-value

frequentist vs Bayesian

Bayes theorem

reproducibility crisis

statistic

parameter

continuous variable

discrete variable

experimental vs observational studies

μ

σ

$\bar{Y}$

s

blinding

pseudoreplication

biological vs technical replicates

outliers

confounding variables

common faults in plots

sample

population

transformation

parametric

non-parametric

MCMC

R skills you should have

Create matrices, vectors, dataframes, and lists

Subset each of these objects

Handle csv files to import/export data

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

Perform a permutation or Monte Carlo test

Perform and correctly interpret the statistical tests mentioned below

Evaluate an MCMC log file

R functions you should handle with ease:

binom.test

chisq.test

t.test (single sample, two sample, paired)

aov

lm

glm

Example Problems

1) Suppose you are studying a pair of cryptic species. In your area 5% of individuals are species A and 95% of individuals are species B. There is currently no genetic assay capable of telling them apart. They differ however in the frequency of a rare color pattern. Species A has the rare color pattern 50% of the time while species B has the rare color pattern only 2% of the time. Assume these numbers are known with certainty, from many years of field research. Now suppose you find one of these species with the rare color pattern. Use Bayes theorem to compute the probability that it is from species A.

2) Download the two mcmc log files from the course website. Choose the MCMC that represents a "good" run? Provide a description of the rate parameter for codon2 and codon3.

3) Grasshoppers recover movement of a leg after nerve damage. Download the grasshopper dataset from the course website. It has four columns that describe range of motion before injury, directly after injury, after a 2-week recovery, and then after crushing the primary nerve a second time. The grasshoppers could recover movement by repair to the crushed nerve if so then crushing a second time should cause them to lose range of motion. However, if they are recovering range of motion by utilizing other nerves serving the legs then the recrushing should have no impact on range of motion. Determine whether the grasshoppers are repairing the damaged nerve or using alternate pathways to recover range of motion.

4) Use the retrogene3.csv file from the course website and test whether chromosome 7 has a different number of parents or daughters relative to your expectations.