

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

p-value	experimental vs observational	common faults in plots
frequentist vs Bayesian	studies	sample
approaches	$\mu$	population
Bayes theorem	$\sigma$	transformation
reproducibility crisis	$\bar{Y}$	parametric
statistic	$s$	non-parametric
parameter	blinding	species as datapoints
continuous variable	pseudoreplication	interaction
discrete variable	biological vs technical replicates	MCMC (I'll cover this on Thursday)
nominal variable	outliers	
ordinal variable	confounding variables	

**R skills you should have**

Create matrices, vectors, dataframes, and lists  
Subset each of these objects  
Read a csv file to import 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

**R functions you should handle with ease:**

binom.test	lm
chisq.test	glm
t.test (single sample, two sample, paired)	prcomp
aov	

**Example Problems**

Download the frog.csv morphology dataset from the website this data includes measures for 3 species for 10 different traits as well as 10 unidentified species. To what species do these 10 unidentified samples belong.

Download the az.csv data from the website this contains occurrence data for Chrysina gloriosa and C. beyeri. What tree do you believe is most important in determining whether each species is present?

Download the retro.copies.csv data from the website this contains chromosome 2 is a new sex chromosome does it exhibit unexpected patterns of retrogene generation or acquisition?

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.