

2024 Final Review

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

p-value	discrete variable	common faults in plots
frequentist vs Bayesian approaches	experimental vs observational studies	sample
Bayes theorem	blinding	population
reproducibility crisis	pseudoreplication	transformation
statistic	biological vs technical replicates	parametric
parameter	outliers	non-parametric
continuous variable	confounding variables	interaction
		MCMC

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	while	plot
chisq.test	for	sum
t.test	polygon	sample
aov	rgb	rnorm
lm	abline	rexp
glm	text	runif
if	points	

Example Problems

Download the az.csv data from the website that contains occurrence data for *Chrysina gloriosa* and *C. beyeri*. Use a GLM to determine what tree you believe is most important in determining whether each species is present?

Download the retro.copies.csv data from the website. In this species 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.

Download the hybridization.csv file from the course website. Use a logistic regression and determine the divergence that is associated with a 50% chance of exhibiting Haldane’s rule and the divergence that is associated with a 50% chance of being “good species” (unable to hybridize).

Make a publication quality plot of an interesting pattern you identify in the oaks.csv data file.

Example Hard Questions

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?

You are interested in determining whether skewed sex ratios (not 50/50) are more common in species with XY or ZW sex chromosomes use the file offspring.csv to determine whether or not skewed sex ratios are more common under one sex chromosome system or the other (lots of ways that you could do this).