

This is a preview of what students will see when they are submitting the assignment. Interactive features are disabled.

Assignment 2

Due: Friday October 2, 2020 11:59 PM (Atlantic Daylight Time)

Assignment description

If you use R in a question, make sure you include the relevant R commands and output. R output can be integrated into your answer or be placed at the end of the question.

R output alone is almost always only **part** of the answer to a question.

Note that in many statistical questions, there may be more than one good answer to a particular question, but that some answers may be better than others.

Submit your assignment

 [Help](#)

After you have completed the assignment, please save, scan, or take photos of your work and upload your files to the questions below. Crowdmark accepts PDF, JPG, and PNG file formats.

Q1 (5 points)

The data in the following link gives the photosynthetic capacity of nine individual females of the neotropical tree *Ocotea tenera*, according to the number of fruits produced in the previous reproductive season (Wheelwright and Logan 2004). The goal of the study was to investigate how reproductive effort in females of these trees impacts subsequent growth and photosynthesis.

(<https://whitlockschluter3e.zoology.ubc.ca/Data/chapter02/chap02q26NeotropicalTreePhotosynthesis.csv>)

(a) Graph the association between these two variables using the most appropriate method. Identify the type of graph you used.

(b) Which variable is the explanatory variable in your graph?

(c) Describe the association between the two variables in words, as revealed by your graph.

Q2 (8 points)

Researchers have created every possible "knockout" line in yeast. Each line has exactly one gene deleted and all the other genes present (Steinmetz et al. 2002). The growth rate -- how fast the number of cells increases per hour -- of each of these yeast lines has also been measured, expressed as a multiple of the growth rate of the wild type that has all the genes present. In other words, a growth rate greater than 1 means that a given knockout line grows faster than the wild type, whereas a growth rate less than 1 means it grows more slowly. Below is the growth rate of a random sample of knockout lines:

0.86, 1.02, 1.02, 1.01, 1.02, 1, 0.99, 1.01, 0.91, 0.83, 1.01

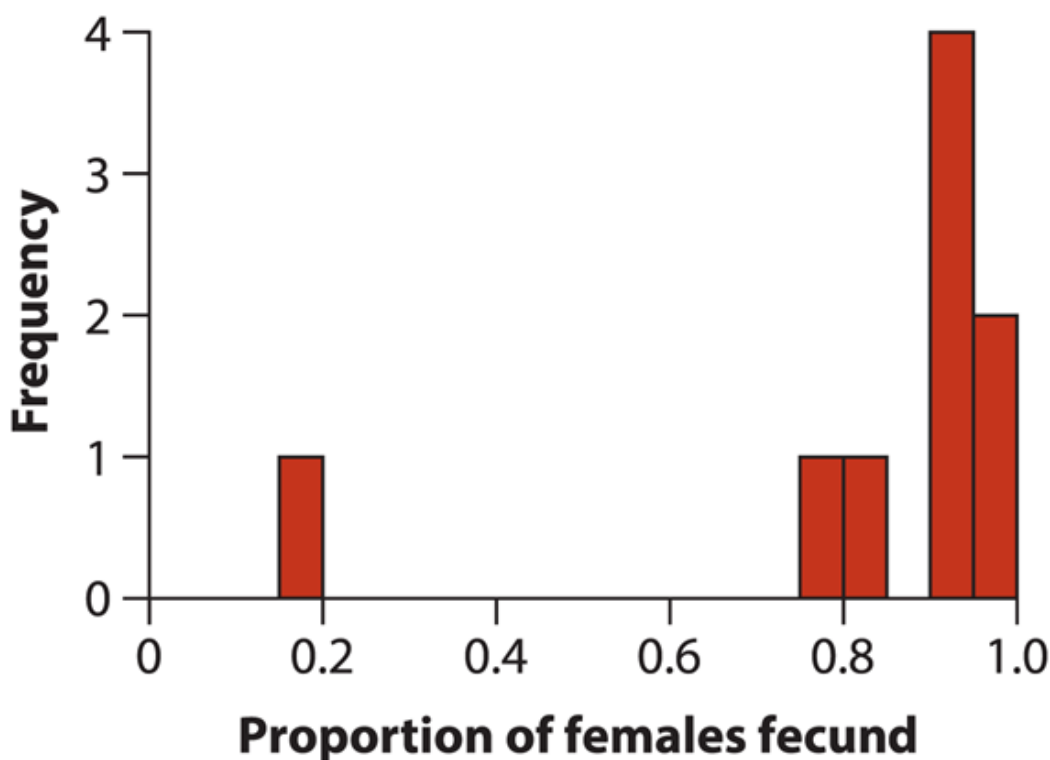
(a) Calculate the mean, variance, standard deviation and coefficient of variation of this sample of yeast lines.

(b) Calculate the five-number summary, i.e., the minimum, lower quartile, median, upper quartile and the maximum of the sample.

(c) Calculate the interquartile range (IQR), the lower and upper fences. Are there any outliers?

Q3 (4 points)

Consider this graph:



(a) What would be the better choice to describe the typical value in the data, the mean or the median? Why?

(b) What would be the better choice to describe the spread of data around their center, the standard deviation or the interquartile range? Why?

Q4 (5 points)

As in other vertebrates, individual zebrafish differ from one another along the shy-bold behavioral spectrum. In addition to other differences, bolder individuals tend to be more aggressive, whereas shy individuals tend to be less aggressive. Norton et al. (2011) compared several behaviors associated with this syndrome between zebrafish that had the *spiegeldanio* (*spd*) mutant at the *Fgf1a* gene and the "wild type" lacking the mutation. The data in the link and below are measurements of the amount of the time, in seconds, that individual zebrafish with and without this mutation spent in aggressive activity over 5 minutes when presented with a mirror image.

(<https://whitlockschluter3e.zoology.ubc.ca/Data/chapter03/chap03q23ZebraFishBoldness.csv>)

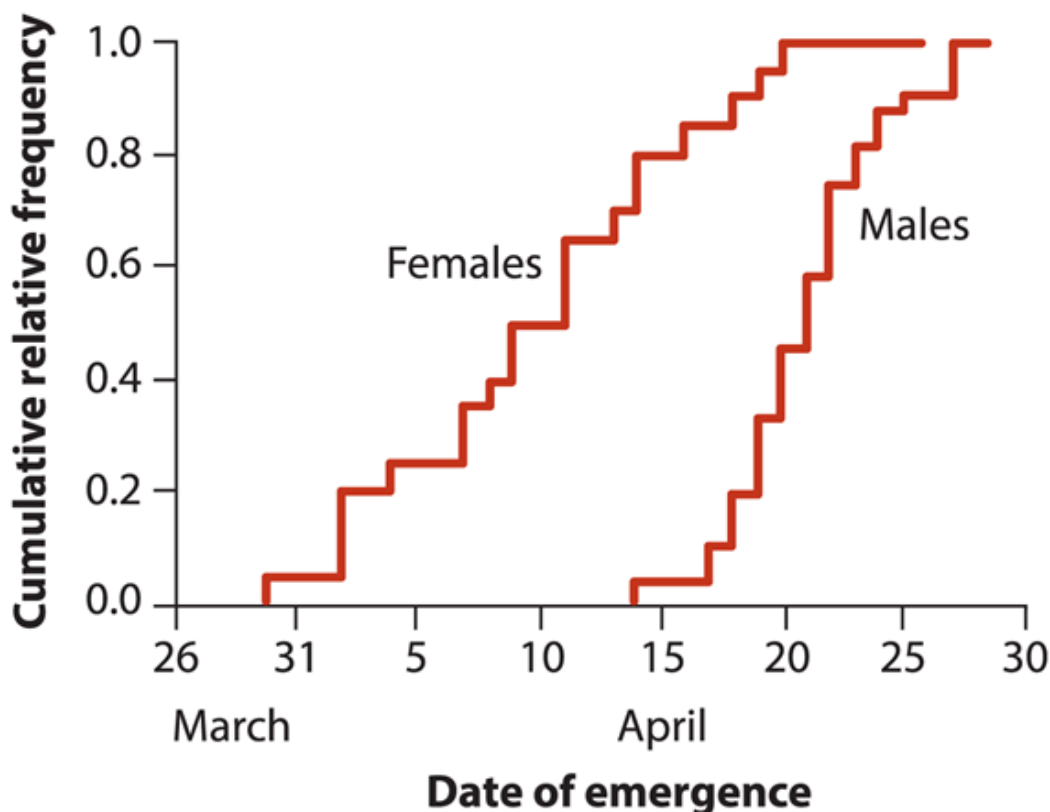
Wild type: 0, 21, 22, 28, 60, 80, 99, 101, 106, 129, 168

Spd mutant: 96, 97, 100, 127, 128, 156, 162, 170, 190, 195

- Draw a (side-by-side) boxplot to compare the frequency distributions of aggression score in the two groups of zebrafish.
- According to the boxplot, which genotype has the higher aggression scores?
- Which sample spans the higher range of values for aggression scores?
- Which sample has the larger interquartile range?

Q5 (2 points)

As the Arctic warms and winters become shorter, hibernation patterns of arctic mammals are expected to change. Sheriff et al. (2011) investigated emergence dates from hibernation of arctic ground squirrels at sites in the Brooks Range of northern Alaska. The measurements shown in the figure below are emergence dates in a sample of male and female ground squirrels at one of their study sites.



- (a) Which sex, females or males, has the earlier median emergence date?
- (b) Which sex, females or males, has the greater interquartile range in emergence date?