Nathan Yee

BIOL 171

Ryan Hill

Lab 11: Working with Data, Excel T-Test And intro to R

Part 1

1. The sample size for male and female *S. callippe* for this dataset is 21 and 16 butterflies respectively.
2. The minimum for males and females is 0.148 g and 0.223 g. The maximum is 0.242 g and 0.323 g.
3. This data seems reasonable. Females are for the most part, larger than their male counterparts.
4. A graph of a person and person

   Description automatically generated
5. There is on outlier, a female butterfly that weighed 0.0223 g. This is significantly smaller than the rest and may be an incorrectly entered value.
6. The average for males and females is 0.188 g and 0.283 g respectively.
7. Yes, the male and female distributions overlap but they look different.

Part 2

A graph of different colored bars

Description automatically generated

1. Our distributions are normally distributed for the most part. The curve for males looks like a normal bell curve histogram, but the female curve is biased to the right.
2. The variance for males is 0.00534 and the variance for females is 0.00136.
3. They do not have equal variances, the data is therefore heterscedastic
4. Females have the higher variance value
5. The difference in variation reflects well in the distributions plotted above. The curve for females is skewed to the right. This makes sense because female data is more spread out.

Our P-value is 2.71 x 10-6

Our t-statistic value is -5.75

1. Based on the t-test with a p-value of 2.70946E-06 and a t-statistic of -5.750985033, we reject the null hypothesis that male and female *Speyeria callippe* have the same body mass distribution. This suggests that there is evidence of sexual dimorphism in body mass, with males and females having statistically significantly different average masses.

III.A

A graph of a person and person

Description automatically generated

A graph of a person with a number of numbers

Description automatically generated with medium confidence

1. The histograms look like they overlap, but only by a bit. Histograms from both programs seem to agree and look similar, however, RStudio’s graphics are accurate.

A diagram of a graph

Description automatically generated

1. The shape of a box plot summarizes the central tendency, spread, skew, and outliers in a dataset, while histograms display the same information in terms of peak locations, width, skew, and range. Histograms are generally more detailed. In a box plot, the central rectangle is the spread of the middle 50% of the data. The line within the box represents the median or the 50th percentile. Outliers are often displayed as points beyond the whiskers. The whiskers of the box plot detail the range and spread of the data. The length of the whiskers is how spread out the data is. If one whisker is longer than the other, it suggests a skew in the data.
2. Variance

Males = 0.0005335588

Females = 0.001362711

Standard Deviation

Males = 0.02309889

Females = 0.03691492

* A larger standard deviation tends to correspond to a wider, more dispersed histogram (females) while a smaller standard deviation tends to correspond to a narrower, more concentrated histogram (males)

1. P-value = 5.419e-06

T-value = 7.3015

* With a p-value of 5.419e-06 and a t-value of 7.3015, there is a significant difference between the two populations.

1. We reject the null hypothesis of males and females coming from the same distribution. There is a statistically significant difference in body mass between male and female populations of *S. callippe*. This is statistically significant evidence for sexual dimorphism.