New York University Tandon School of Engineering

Biomedical Engineering
Applied Mathematics and Statistics for Biomedical Engineering

Fall 2021
Professor Mirella Altoe
Tuesday 5:00-7:30PM Rogers Hall 325

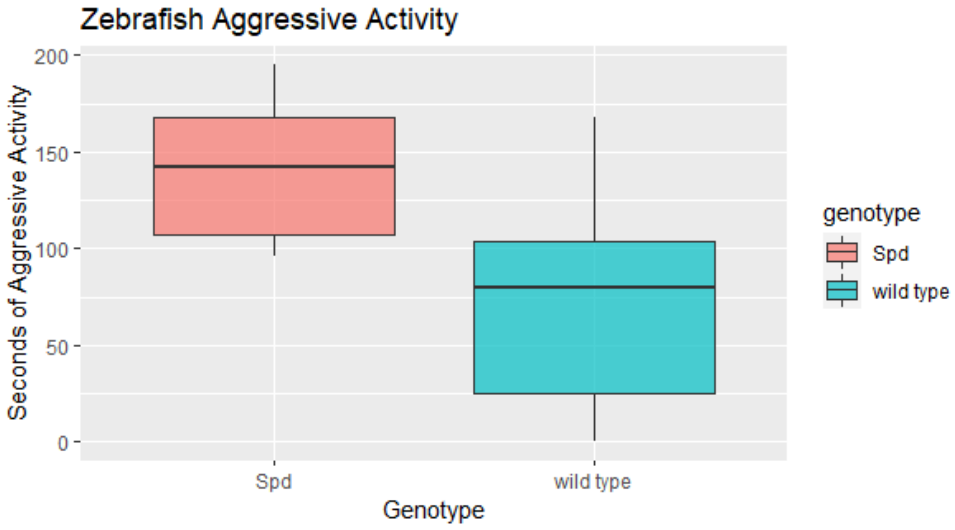
Computer Lab Assignment #2

QUESTION 1: 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 *Fgfr1a* gene (reduced fibroblast growth factor receptor 1a) and the "wild type" lacking the mutation. The data below are measurements of the amount of time, in seconds, that individual zebrafish with and without this mutation spent in aggressive activity over 5 minutes when presented with a mirror image. Data available here.

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

A. **Draw a boxplot to compare the frequency distributions of aggression score in the two groups of zebrafish. According to the box plot, which genotype has the higher aggression scores?**
According to the boxplot created in R, the Spd mutant zebrafish has higher aggression scores.



B. **According to the box plot, which sample spans the higher range of values for aggression scores?**
The wild type zebrafish has the higher range of aggression scores (as shown by the whiskers on the box plot).

C. **Which sample has the larger interquartile range?**
The wild type zebrafish has a larger interquartile range (of 78.5, compared to the mutant's 61.25). This is indicated by the "box" portion of each plot.

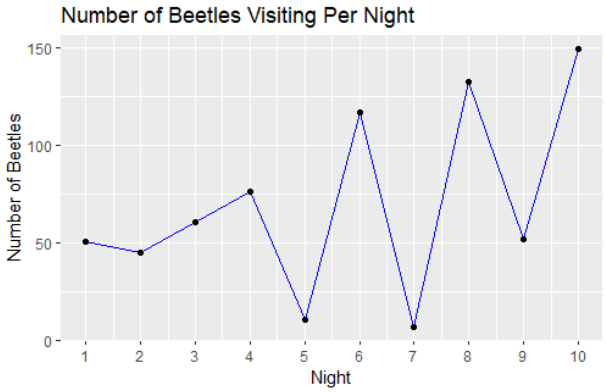
D. **What are the vertical lines projecting outward above and below each box?**
The vertical lines projecting outward above and below indicates the upper and lower non-extreme values in the dataset (extreme values are defined as: (Q1-1.5*IQR) or (Q3+1.5*IQR).

QUESTION 2: *Amorphophallus johnsonii* is a plant growing in West Africa, and it is better known as a "corpse flower." Its common name comes from the fact that when it flowers, it gives off a "powerful aroma of rotting fish and faeces" (Beath 1996). The flowers smell this way because their principal pollinators are carrion beetles, who are attracted to such a smell. Beath (1996) observed the number of carrion beetles (*Phaeochrous amplus*) that arrive per night to flowers of this species. Data available here.

A. **What is the mean and standard deviation of beetles per flower?**
The mean is 70.1 and the standard deviation is 48.5.

B. **What is the standard error of this estimate of the mean?**
The standard error is 15.33.

C. **Redraw the figure using the most appropriate method discussed in this chapter. What type of graph did you use? Give an approximate 95% confidence interval of the mean. Provide lower and upper limits.**
Since the data is number of carrion beetles visiting Phaeochrous amplus per night, I used a line graph. The approximate 95% confidence interval is calculated by the mean +/- (2SE). Thus, the lower and upper bound is 39.44 and 100.76 respectively.



D. **If you had been given 25 data points instead of 10, would you have expected the standard error of the mean to be greater than, less than, or about the same as this sample?**
Since standard error is calculated by the standard deviation divided by the square root of the sample size, the standard error would be lower given a larger sample size.