Simonson_HW4

Marty Simonson

February 14, 2019

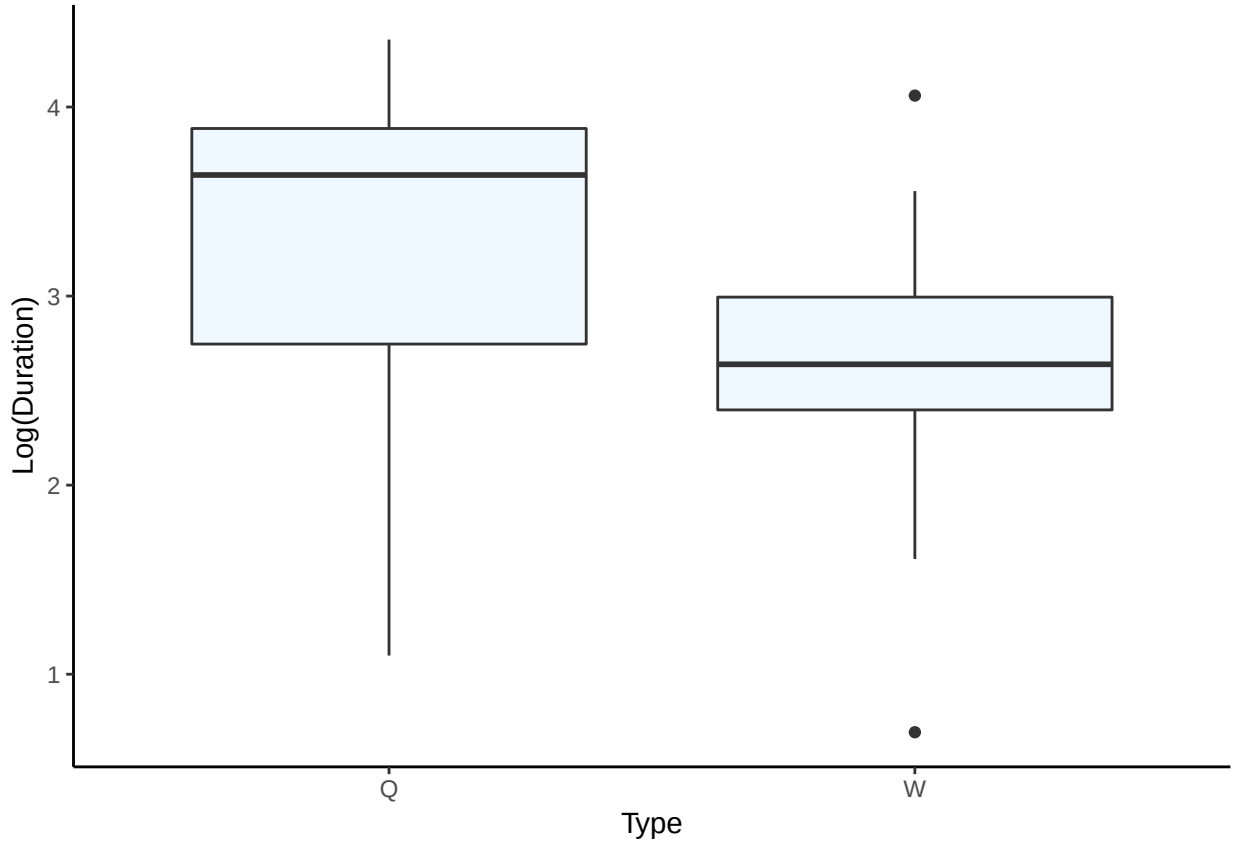
1.) Pollen Removal

Using data from *bee.csv*, test if log-transformation helps meet assumptions of pooled t-tests.

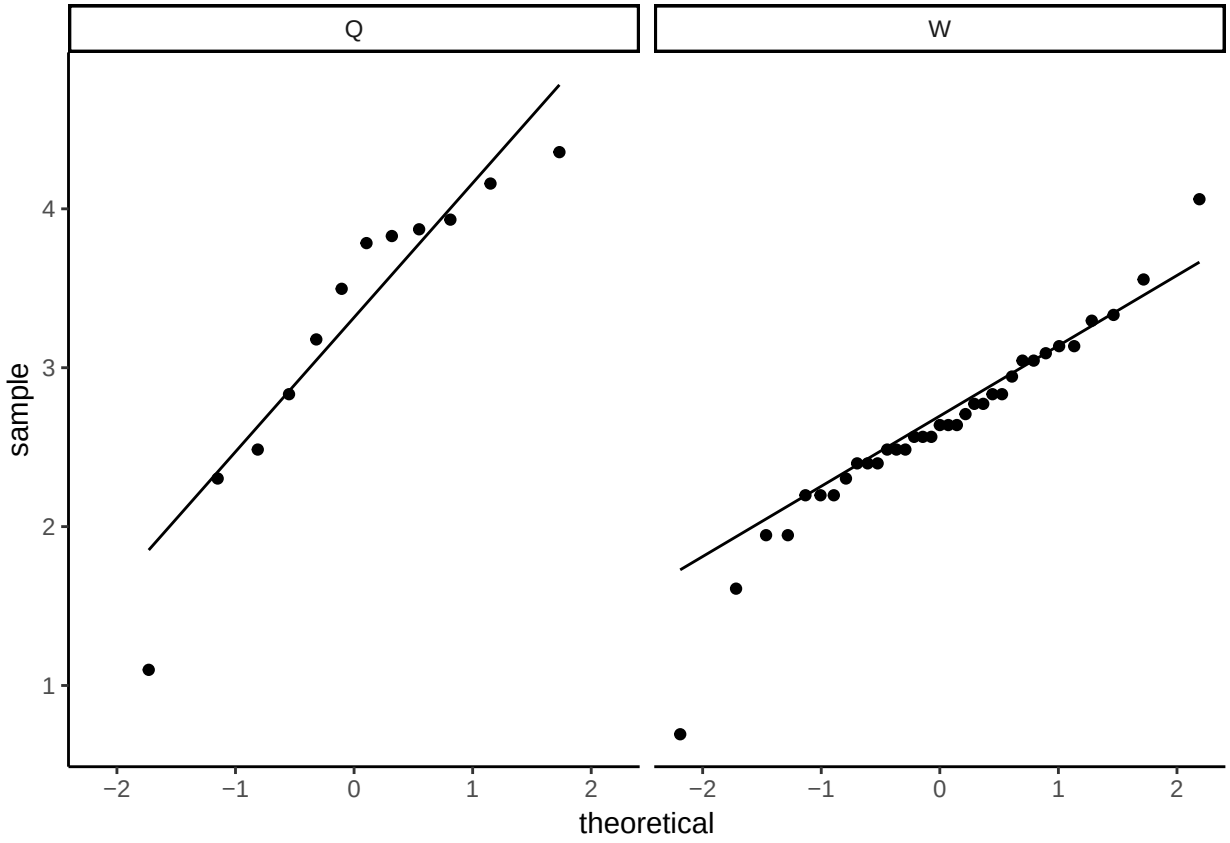
a) Construct box plots and normal quantile plots on the log-transformed duration values for both types of bees.

```
df<-read.csv("bee.csv")
df$LogDuration<-log(df$Duration)

ggplot(data=df,aes(x=Type, y=LogDuration))+
       geom_boxplot(fill="aliceblue")+
       ylab("Log(Duration)")+
       theme_classic()
```




```
ggplot(data=df, aes(sample=LogDuration)) +
        stat_qq() +
        stat_qq_line() +
        facet_wrap(facets = vars(Type))+
        theme_classic()
```



Does it appear that assuming normality for thelog-transformed duration variable is reasonable? Explain. - **Answer:** The assumption of normality is not met after log-transformation because there is still curvature in the points along the QQ line.

Are the variabilities the same for the two types of bees on the log scale? Explain. - **Answer:** The size of the boxes and whiskers are not equal, therefore the assumption of normality is not met.

b) Use the Welch's t-test to compare the (population) median durations. Write down the3 null and alternate hypotheses, provide the t-stastic and p-value.