Course Project pt2

Zach Dungan

3/3/2022

## Part 2: Inferential Data Analysis

Load libraries and ToothGrowth data

library(ggplot2)

## Warning in register(): Can't find generic `scale\_type` in package ggplot2 to  
## register S3 method.

library(datasets)  
  
data("ToothGrowth")

**View info about ToothGrowth data**

str(ToothGrowth)

## 'data.frame': 60 obs. of 3 variables:  
## $ len : num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...  
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...  
## $ dose: num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...

head(ToothGrowth)

## len supp dose  
## 1 4.2 VC 0.5  
## 2 11.5 VC 0.5  
## 3 7.3 VC 0.5  
## 4 5.8 VC 0.5  
## 5 6.4 VC 0.5  
## 6 10.0 VC 0.5

Summary statistics

summary(ToothGrowth)

## len supp dose   
## Min. : 4.20 OJ:30 Min. :0.500   
## 1st Qu.:13.07 VC:30 1st Qu.:0.500   
## Median :19.25 Median :1.000   
## Mean :18.81 Mean :1.167   
## 3rd Qu.:25.27 3rd Qu.:2.000   
## Max. :33.90 Max. :2.000

Correct ‘dose’ varibale

ToothGrowth$dose <- as.factor(ToothGrowth$dose)

**Data Analysis**

supp\_mean <- split(ToothGrowth$len, ToothGrowth$supp)  
sapply(supp\_mean, mean)

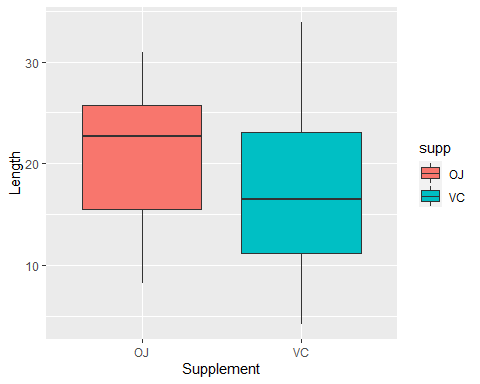
## OJ VC   
## 20.66333 16.96333

supp\_dose <- split(ToothGrowth$len, ToothGrowth$dose)  
sapply(supp\_dose, mean)

## 0.5 1 2   
## 10.605 19.735 26.100

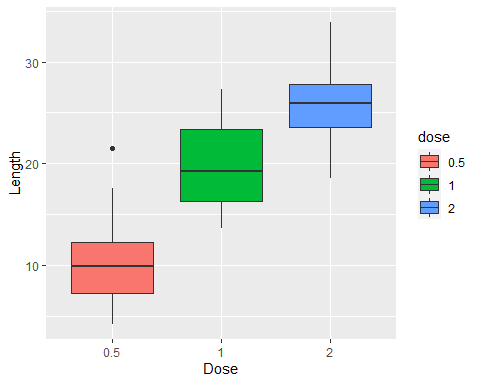
Graph of effect of supplement on tooth length:

ggplot(aes(x=supp, y=len), data=ToothGrowth) + geom\_boxplot(aes(fill=supp)) + xlab("Supplement") + ylab("Length")



Graph of effect of supplement on tooth length:

ggplot(aes(x=dose, y=len), data=ToothGrowth) + geom\_boxplot(aes(fill=dose)) + xlab("Dose") + ylab("Length")



**Statistical Tests**

Create easy to use variables

supp <- ToothGrowth$supp  
dose <- ToothGrowth$dose  
len <- ToothGrowth$len

T-Test #1

# Check for differences in supplement type  
t.test(len[supp == "VC"], len[supp == "OJ"])

##   
## Welch Two Sample t-test  
##   
## data: len[supp == "VC"] and len[supp == "OJ"]  
## t = -1.9153, df = 55.309, p-value = 0.06063  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -7.5710156 0.1710156  
## sample estimates:  
## mean of x mean of y   
## 16.96333 20.66333

The p-value is about 0.06, meaning we cannot reject the null hypothesis that different supplements don’t have an effect on tooth length.

T-Test #2

# Check for differences in dosage type  
# In this test, use 0.5 and 1.0  
t.test(len[dose == 0.5], len[dose == 1.0])

##   
## Welch Two Sample t-test  
##   
## data: len[dose == 0.5] and len[dose == 1]  
## t = -6.4766, df = 37.986, p-value = 1.268e-07  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -11.983781 -6.276219  
## sample estimates:  
## mean of x mean of y   
## 10.605 19.735

The p-value is very close to zero, which shows a considerable impact on tooth length, meaning we can reject the null hypothesis.

For thoroughness, the tests between [0.5 and 2.0] and [1.0 and 2.0] are below, with similar results

# 0.5 and 2.0  
t.test(len[dose == 0.5], len[dose == 2.0])

##   
## Welch Two Sample t-test  
##   
## data: len[dose == 0.5] and len[dose == 2]  
## t = -11.799, df = 36.883, p-value = 4.398e-14  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -18.15617 -12.83383  
## sample estimates:  
## mean of x mean of y   
## 10.605 26.100

# 1.0 and 2.0  
t.test(len[dose == 1.0], len[dose == 2.0])

##   
## Welch Two Sample t-test  
##   
## data: len[dose == 1] and len[dose == 2]  
## t = -4.9005, df = 37.101, p-value = 1.906e-05  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -8.996481 -3.733519  
## sample estimates:  
## mean of x mean of y   
## 19.735 26.100

**Conclusions**

1. The supplement type, statistically speaking, has no significant effect on tooth length.
2. The dosage amount has a significant effect on tooth length