Inferential Data Analysis

Alberto Macías

9/29/2020

Synopsis

We realize and inferential data analysis on the ToothGrowth data set in order to study how vitamin C provided via orange juice or via ascorbic acid affects in the tooth length in 60 guinea pigs.

Loading the data and performing the EDA

```
library(ggplot2)
library(dplyr)
```

```
library(datasets)
data("ToothGrowth")
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 ...
```

According to R documentation, Toothgrowth data set contains data about the "length of odontoblast (which are cells responsible for tooth growth) in 60 guinea pigs". It is a data frame with 3 variables and 60 observations, one for each guinea pig. The variables are len, a numeric vector that contains the lengths of the cells in unspecified units; supp, a two level factor that indicates the supplement type of vitamin C: orange juice (OJ) or ascorbic acid (VC); and dose, a numeric vector that contains the dosage in which vitamin C was provided, in 3 dosages, 0.5, 1 and 2 milligrams per day (mg/day).

We divide ToothGrowth data set into to data frames, vc and oj, divided by the supplement type. We compute mean, standard deviation and variance of the cell lengths

```
vc <- subset(ToothGrowth, supp=="VC")
oj <- subset(ToothGrowth, supp=="OJ")
mnvc <- mean(vc$len); varvc <- var(vc$len)
mnoj <- mean(oj$len); varoj <- var(oj$len)
mnvc; varvc;
```

```
## [1] 16.96333
```

```
## [1] 68.32723
```

```
mnoj; varoj
```

```
## [1] 20.66333
```

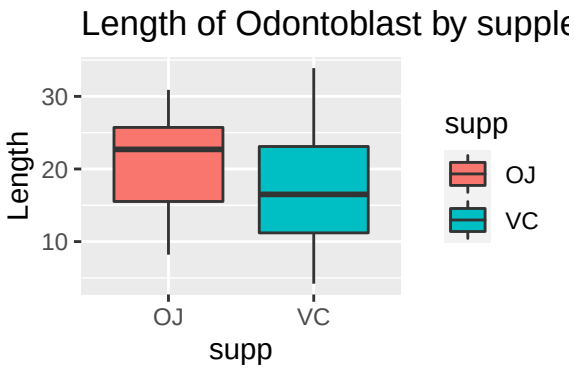
```
## [1] 43.63344
```

We note that the mean of length of cells in oj data is greater than mean of length in vc, which could indicate that the orange juice supplement of vitamin c affects more to cell length than ascorbic acid. Also that data

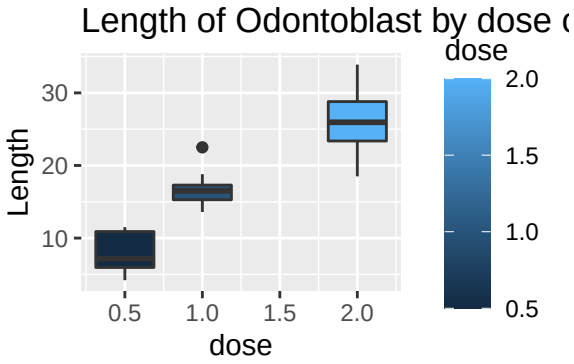
in the oranjuice set is less variable.

Appendix

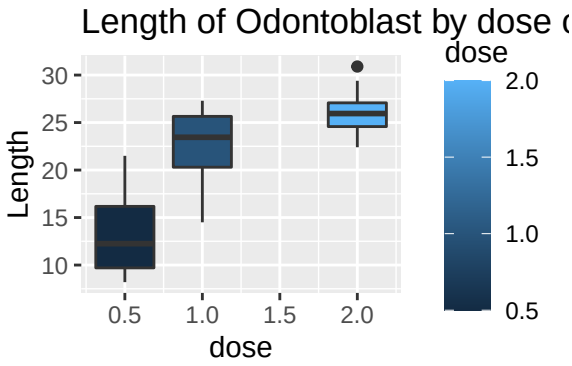
```
ggplot(ToothGrowth, aes(x=supp, y=len, fill= supp))+
            geom_boxplot()+
            ggtitle("Length of Odontoblast by supplement")+
            ylab("Length")
```




```
ggplot(vc, aes(x=dose, y=len, fill=dose, group = dose))+
            geom_boxplot()+
            ggtitle("Length of Odontoblast by dose of ascorbic acid.")+
            ylab("Length")
```




```
ggplot(oj, aes(x=dose, y=len, fill=dose, group = dose))+
            geom_boxplot()+
            ggtitle("Length of Odontoblast by dose of orange juice.")+
            ylab("Length")
```


Length of Odontoblast by dose
Length
30
25
20
15
10
0.5
1.0
1.5
2.0
dose
dose
2.0
1.5
1.0
0.5