Tooth Growth Data Analysis

Loc Nguyen

9/24/2021

Overview

- Load the ToothGrowth data and perform some basic exploratory data analyses
- Analyze the ToothGrowth data by summary, visualization, statistics,...

Loading libraries

Loading necessary libraries for the report

```
library(tidyverse)
library(ggplot2)
library(datasets)
```

Load the ToothGrowth data and perform some basic exploratory data analyses

Load data

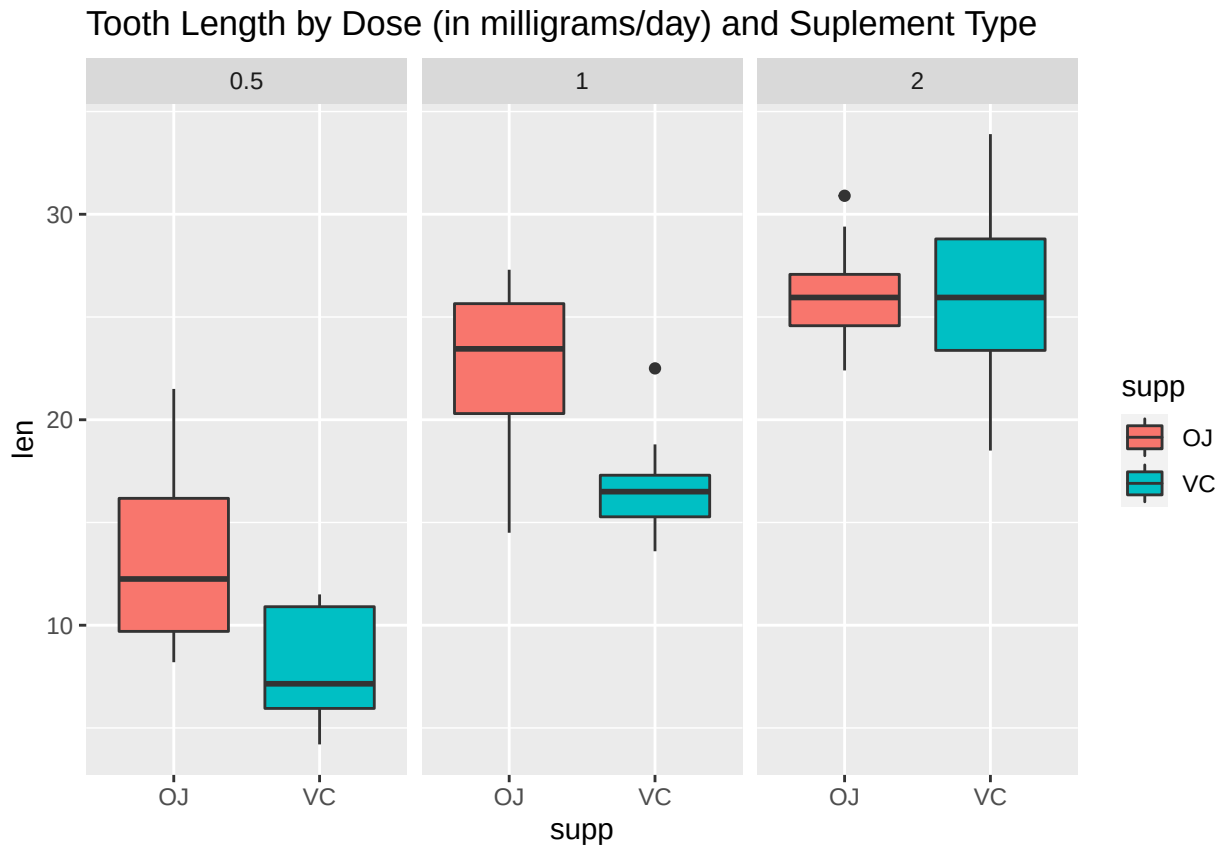
```
data(ToothGrowth)
head(ToothGrowth)
```

```
##    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
```

Perform some basic exploratory data analyses

Comparing the data sets. Some plots to explore data

```
ggplot(data=ToothGrowth, aes(x= supp, y = len,fill = supp)) +
      geom_boxplot() +
      facet_grid(.~dose) +
      ggtitle("Tooth Length by Dose (in milligrams/day) and Suplement Type")
```



Provide a basic summary of the data.

Number of observations for each treatment/experiment. Total guinea pigs = 60

```
table(ToothGrowth$supp, ToothGrowth$dose)
```

```
##
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

Summarize 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 ...
```

```
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
```

Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.

The data contains: tooth length, supplement, and supplement dose for a group of guinea pigs. Presumably the data was not collected in a pairwise fashion. Performing T Test for each subset of data separated by dosages where the null hypothesis is that the tooth growth means of the two supp types

Assumptions

- Samples are both unpaired and unequal variances
- The sample population distribution is mound shaped and not skewed

Test

Subset data by each value of dosage

```
d.5 <- filter(ToothGrowth, dose == 0.5)
d1 <- filter(ToothGrowth, dose == 1)
d2 <- filter(ToothGrowth, dose == 2)
```

Test for dosage = 0.5

```
test.5 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d.5)
test.5
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means between group OJ and group VC is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

Test for dosage = 1

```
test1 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d1)
test1
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means between group OJ and group VC is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Test for dosage = 2

```
test2 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d2)
test2
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means between group OJ and group VC is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Conclusions

P-values are lower than 0.05 for dose 0.5 (0.0063586) and 1.0 (0.0010384) but greater than for 2.0 (0.9638516).

We can say the supplements have different effects for dosages 0.5 and 1.0 but there is not enough evidence to prove so when the dosage is 2.0.