

Part2 Analyzing ToothGrowth

By Zoey Phay

Overview

For the second part, we're going to analyze the ToothGrowth data in the R datasets package.

The dataset records Vitamin C effects on the tooth growth on guinea pigs at different dosage (0.5, 1 and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

A data frame with 60 observations on 3 variables.

len numeric Tooth length supp factor Supplement type (VC or OJ). dose numeric Dose in milligrams.

Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
library(datasets)
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
```

```
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 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Provide a basic summary of the data.

```
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
table(ToothGrowth$supp, ToothGrowth$dose)
```

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

```
summary(ToothGrowth)
```

```
##      len      supp      dose
## Min.   : 4.20   OJ:30   0.5:20
## 1st Qu.:13.07   VC:30   1  :20
## Median :19.25           2  :20
## Mean   :18.81
## 3rd Qu.:25.27
## Max.   :33.90
```

```
mean(ToothGrowth$len) # mean
```

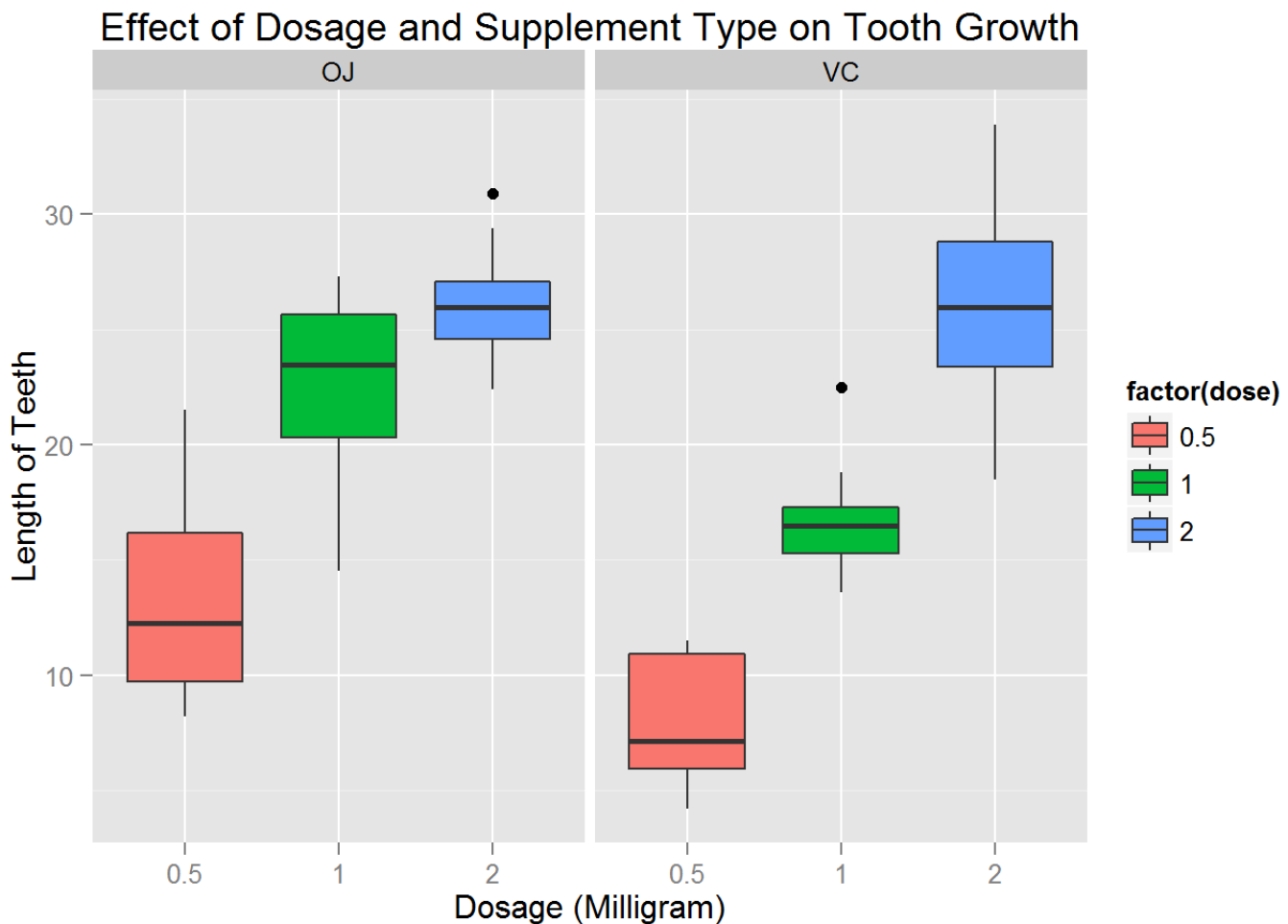
```
## [1] 18.81333
```

```
sd(ToothGrowth$len) # standard deviation
```

```
## [1] 7.649315
```

Graph it..

```
ggplot(ToothGrowth, aes(x=factor(dose),y=len,fill=factor(dose))) + geom_boxplot(notch=F) + facet_grid(
  d(.~supp) +
    scale_x_discrete("Dosage (Milligram)") +
    scale_y_continuous("Length of Teeth") +
    ggtitle("Effect of Dosage and Supplement Type on Tooth Growth")
```



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

```

supp.t1 <- t.test(len~supp, paired=F, var.equal=T, data=ToothGrowth)
supp.t2 <- t.test(len~supp, paired=F, var.equal=F, data=ToothGrowth)
supp.result <- data.frame("p-value"=c(supp.t1$p.value, supp.t2$p.value),
                          "Conf-Low"=c(supp.t1$conf[1],supp.t2$conf[1]),
                          "Conf-High"=c(supp.t1$conf[2],supp.t2$conf[2]),
                          row.names=c("Equal Var","Unequal Var"))

supp.result

```

```

##           p.value  Conf.Low Conf.High
## Equal Var  0.06039337 -0.1670064  7.567006
## Unequal Var 0.06063451 -0.1710156  7.571016

```

State conclusions and the assumptions needed.

Assumptions: Guinea pigs were randomly chosen (same population variance); Samples are independent (sample data not paired)

Conclusion: 1. Orange juice results more tooth growth than ascorbic acid, at 0.5 and 1 dosages. 2. Tooth growth is not significantly different for the different delivery method at 2 mg. 3. Dosage plays the key role in tooth growth