Statistical Inference Project Part 2

Jing Wei Chan

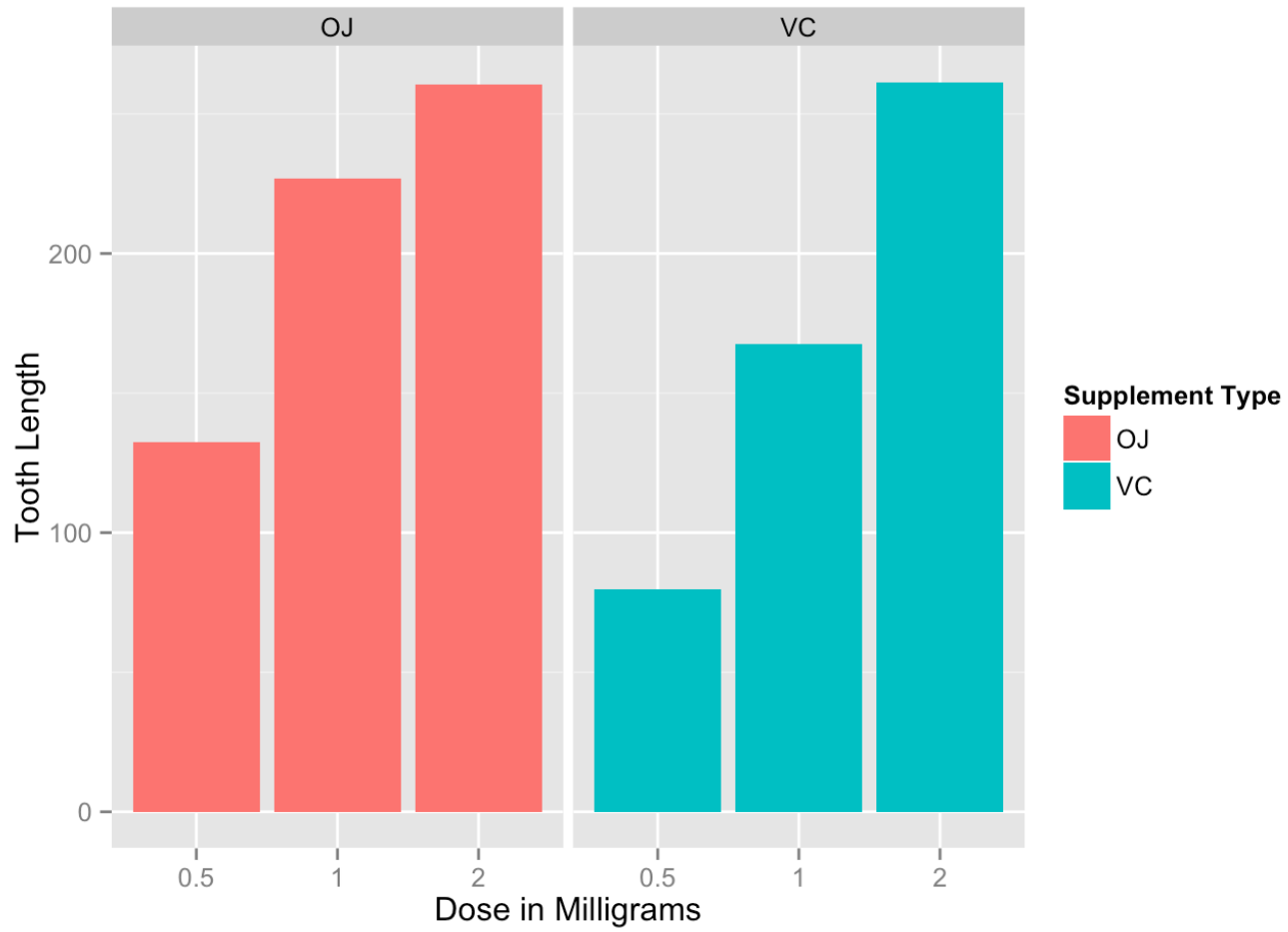
24 August 2015

Overview

In the second portion of the project, we are going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses.

```
library(datasets)
toothgrowth <- ToothGrowth
library(ggplot2)
ggplot(data = toothgrowth, aes(x = as.factor(dose), y = len, fill = supp)) +
    geom_bar(stat = "identity",) +
    facet_grid(. ~ supp) +
    xlab("Dose in Milligrams") +
    ylab("Tooth Length") +
    guides(fill = guide_legend(title="Supplement Type"))
```



As shown, there is a positive relationship between dose and tooth length for both OJ and VC supplement types.

2. Provide a basic summary of the data.

```
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
```

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

We will analyze tooth growth by carrying out t-tests at 95% confidence intervals for the 2 supplements by dosage level.

```
t.test(len ~ supp, toothgrowth[toothgrowth$dose == 0.5, ])
```

```
##
##  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 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
```

At a 0.5 milligram dose, the null hypothesis is rejected as zero does not fall within the confidence interval and p-value < 0.05.

```
t.test(len ~ supp, toothgrowth[toothgrowth$dose == 1, ])
```

```
##
##  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 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
```

At a 1 milligram dose, the null hypothesis is rejected as zero does not fall within the confidence interval and p-value < 0.05.

```
t.test(len ~ supp, toothgrowth[toothgrowth$dose == 2, ])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means 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
```

At a 2 milligrams dose, the null hypothesis is not rejected as zero falls within the confidence interval and p-value > 0.05.

4. State your conclusions and the assumptions needed for your conclusions.

From the exploratory analysis, we can see that there is a positive relationship between dose and tooth length for both OJ and VC supplement types.

From the t-tests, we can conclude that for 0.5 milligram and 1 milligram, OJ is more effective than VC in promoting tooth growth. However, for 1 milligram dose, we are unable to conclude that OJ is more effective than VC.