Statistical Inference Project 2

Scot Shields

Sunday, February 22, 2015

Overview:

We're going to analyze the ToothGrowth data in the R datasets package with summary statistics, plots and confidence intervals.

Load Data & Libraries:

```
cache = TRUE
library(datasets)
library(ggplot2)
```

Summary Statistics:

```
str(ToothGrowth)
```

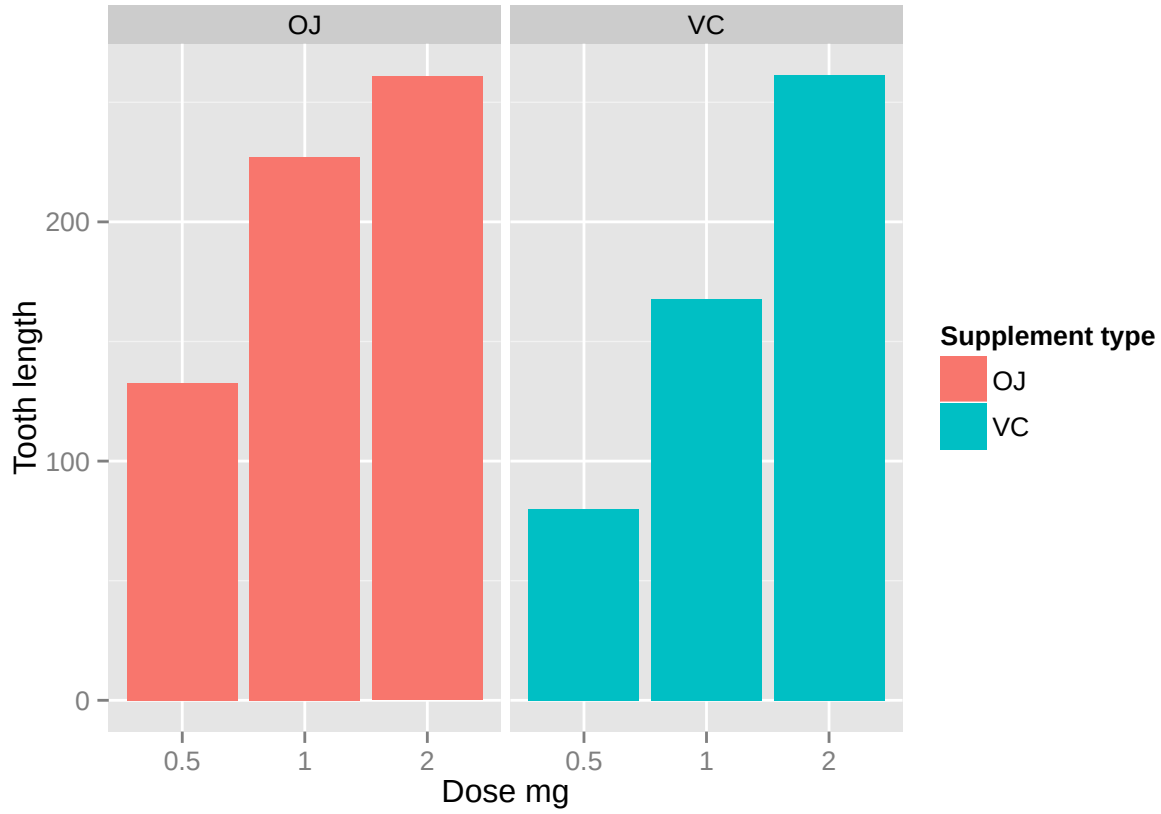
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

Plots of Data:

```
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
    geom_bar(stat="identity",) +
    facet_grid(. ~ supp) +
    xlab("Dose mg") +
    ylab("Tooth length") +
    guides(fill=guide_legend(title="Supplement type"))
```



From the plots we can see a positive correlation between dose amount and tooth length.

Confidence Interval and Hypothesis Test:

Assumptions:

1. The experiment was done with random assignment to different dose level categories and supplement type.
2. Members of the sample population, are representative of the entire population.
3. For the t-tests, the variances are assumed to be different for the two groups being compared.

```
t.test(len ~ supp, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Conclusion:

The p-value is 0.06, and the confidence interval contains zero. This indicates that we can not reject the null hypothesis that the different supplement types have no effect on tooth length.