Inferential Data Analysis on ToothGrowth Data

Shang Ju

January 22, 2017

Load the ToothGrowth Data

```
library(datasets)
data(ToothGrowth)
```

ToothGrowth Dataset Overview

The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).

Summary

```
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
```
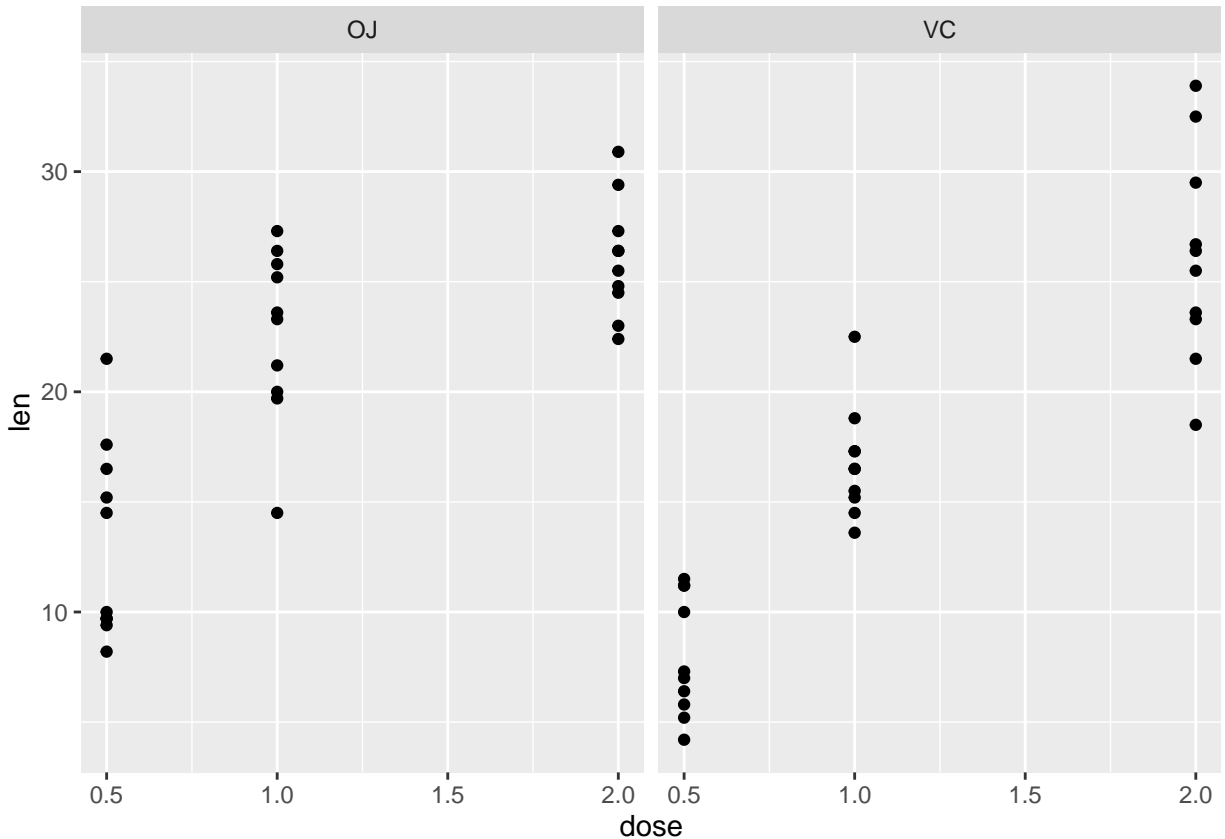
Exploratory Data Analysis

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

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

The dataset is comprised of 60 observations evenly distributed into 6 groups (2 supplements * 3 dose amount)

```
library(ggplot2)
ggplot(data = ToothGrowth,aes(x=dose,y=len)) +
    facet_grid(.~supp) +
    geom_point()
```

From the graph we see that the more dose used, the longer the length of odontoblasts. However, the effectiveness between orange juice (OJ) and Vitamin C (VC) is not significantly different and need further statistical analysis.

Hypothesis Test 1: Tooth Length Affected by Delivery Method (OJ v.s. VC)

```r
t.test(len~supp, data = ToothGrowth, var.equal=F)
```

```
##
##  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
```

Because p-value = 0.06 > 0.05, we can not reject the null hypothesis. So the difference between the effect on tooth growth by delivery methods (orange juice v.s. ascorbic acid) is **not statistically significant**.

Hypothesis Test 2: Tooth Length Affected by Dose Amount

```
t.test(len~dose, data = subset(ToothGrowth,dose %in% c(0.5,1)), var.equal=F)
```

```
##
##   Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

p-value 1.3e-7 << 0.05: the effect on tooth growth by dose amount, 0.5 mg/day v.s. 1 mg/day, is significantly different.

```
t.test(len~dose, data = subset(ToothGrowth,dose %in% c(1,2)), var.equal=F)
```

```
##
##   Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

p-value 1.9e-5 << 0.05: the effect on tooth growth by dose amount, 1 mg/day v.s. 2 mg/day, is significantly different.

Conclusions

Dose amount affect tooth growth significantly: the higher the amount is (between 0.5 and 2 mg/day), the longer the tooth length.

Supplement Type (orange juice v.s. ascorbic acid), however, does **not** significantly affect tooth growth

The conclusions are based on the assumptions that we take 5% as the significance level and that no other factors are affecting the experiment.