Statistical Inference Course Project

Andrew E. Davidson

February 22, 2015

My R Markdown file borrows heavily from
https://github.com/bcaffo/courses/blob/master/06_StatisticalInference/03_01_TwoGroupIntervals
(https://github.com/bcaffo/courses/blob/master/06_StatisticalInference/03_01_TwoGroupIntervals)

---

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

In RStudio run help(ToothGrowth) to see a description of the data set.

Description

The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

```
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 ...
```

```
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
```

```
nrow(ToothGrowth)
```

```
## [1] 60
```

```
nrow(na.omit(ToothGrowth) )
```

```
## [1] 60
```

2. Provide a basic summary of the data.

```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.2   OJ:30   Min.   :0.50
##  1st Qu.:13.1   VC:30   1st Qu.:0.50
##  Median :19.2           Median :1.00
##  Mean   :18.8           Mean   :1.17
##  3rd Qu.:25.3           3rd Qu.:2.00
##  Max.   :33.9           Max.   :2.00
```

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

We'll use "Two Group intervals Statistical Inference"

First create 6 list, one for each combination of treament and dose

```
vc05 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 0.5]
vc1 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 1.0]
vc2 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 2.0]

oj05 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5]
oj1 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1.0]
oj2 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2.0]
```

Now for each does we'll compare the VC vs OJ

```
mean(vc05) - mean(oj05)
```

```
## [1] -5.25
```

```
t.test(vc05, oj05, paired = FALSE, var.equal = TRUE)$conf
```

```
## [1] -8.73 -1.77
## attr(,"conf.level")
## [1] 0.95
```

```
mean(vc1) - mean(oj1)
```

```
## [1] -5.93
```

```
t.test(vc1, oj1, paired = FALSE, var.equal = TRUE)$conf
```

```
## [1] -9.019 -2.841
## attr(,"conf.level")
## [1] 0.95
```

```
mean(vc2) - mean(oj2)
```

```
## [1] 0.08
```

```
t.test(vc2, oj2, paired = FALSE, var.equal = TRUE)$conf
```

```
## [1] -3.563  3.723
## attr(,"conf.level")
## [1] 0.95
```

Lets see if we find a differece between treaments when we ignore the dose

```
vc <- ToothGrowth$len[ToothGrowth$supp == "VC"]
oj <- ToothGrowth$len[ToothGrowth$supp == "OJ"]

mean(vc) - mean(oj)
```

```
## [1] -3.7
```

```
t.test(vc, oj, paired = FALSE, var.equal = TRUE)$conf
```

```
## [1] -7.567  0.167
## attr(,"conf.level")
## [1] 0.95
```

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

regardless of the size of the dose guinea pigs recevied, there does not appear to be a statistically signifigant difference in the amount of tooth growth.

The only assumption I made was groups where independent so that we could not use a paired t test