Hypothesis Testing on ToothGrowth Data

Statistical Inference Project Part II, Class 6 in data science series

Ann Crawford

September 7, 2017

Dependencies

```
##install.packages("dplyr")
```

Tooth Growth Data

The standard R data set, ToothGrowth, measures the effect of Vitamin C on tooth growth in guinea pigs 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(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
```

Hypotesis Testing

The sample data graph (see Appednix) implies two Hypothesis: 1. Supply using orange juice as the delivery method induces longer tooth growth. 2. Increase of vitamin C does induces longer tooth growth. — ## Testing the affect of supp, delivery method **HO = mean OJ = mean VC** Divide the data into two groups, OJ and VC. And run an unpaired t test on the groups. T

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.3.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(magrittr) # for %$% extraction

OJ <- ToothGrowth %>% filter(supp=="OJ")
```

```
## Warning: package 'bindrcpp' was built under R version 3.3.3
```

```
VC <- ToothGrowth %>% filter(supp=="VC")

t.test(OJ$len, VC$len)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ$len and VC$len
## 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 of x mean of y
##  20.66333  16.96333
```

There are 3 groups of doses, 0.5, 1, 2

Divide the data into 3 groups, OJ and VC and run t test.

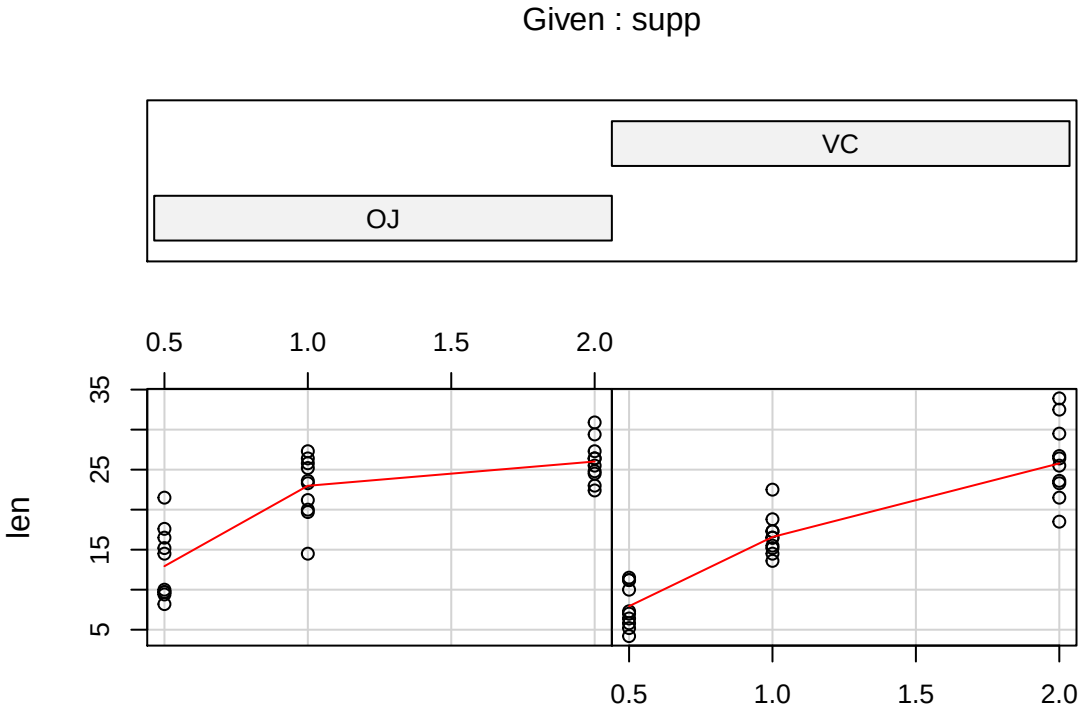
```
low     <- ToothGrowth %>% filter(dose == 0.5)
medium  <- ToothGrowth %>% filter(dose == 1)
high    <- ToothGrowth %>% filter(dose == 2)

t.test(low$len, high$len)
```

```
##
##  Welch Two Sample t-test
##
## data:  low$len and high$len
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

```
t.test(medium$len, high$len)
```

```
##
##  Welch Two Sample t-test
##
## data:  medium$len and high$len
## 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 of x mean of y
##    19.735    26.100
```

Appendix



ToothGrowth data: length vs dose, given type of supplement