Statistical Inference: ToothGrowth

Francisco Marco-Serrano

Summary

According to R Help, the ToothGrowth dataset contains the information on the length of 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). We analyse the differences in length growth by dose level and delivery method.

Data Exploratory Analysis

```
library(datasets)
library(lattice)

data.fr <- ToothGrowth

# Basic information on dimension of the data frame and the variables
dim(data.fr)
```

```
## [1] 60  3
```

```
str(data.fr)
```

```
## '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 ...
```

```
# including summary
summary(data.fr)
```

```
##       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
```

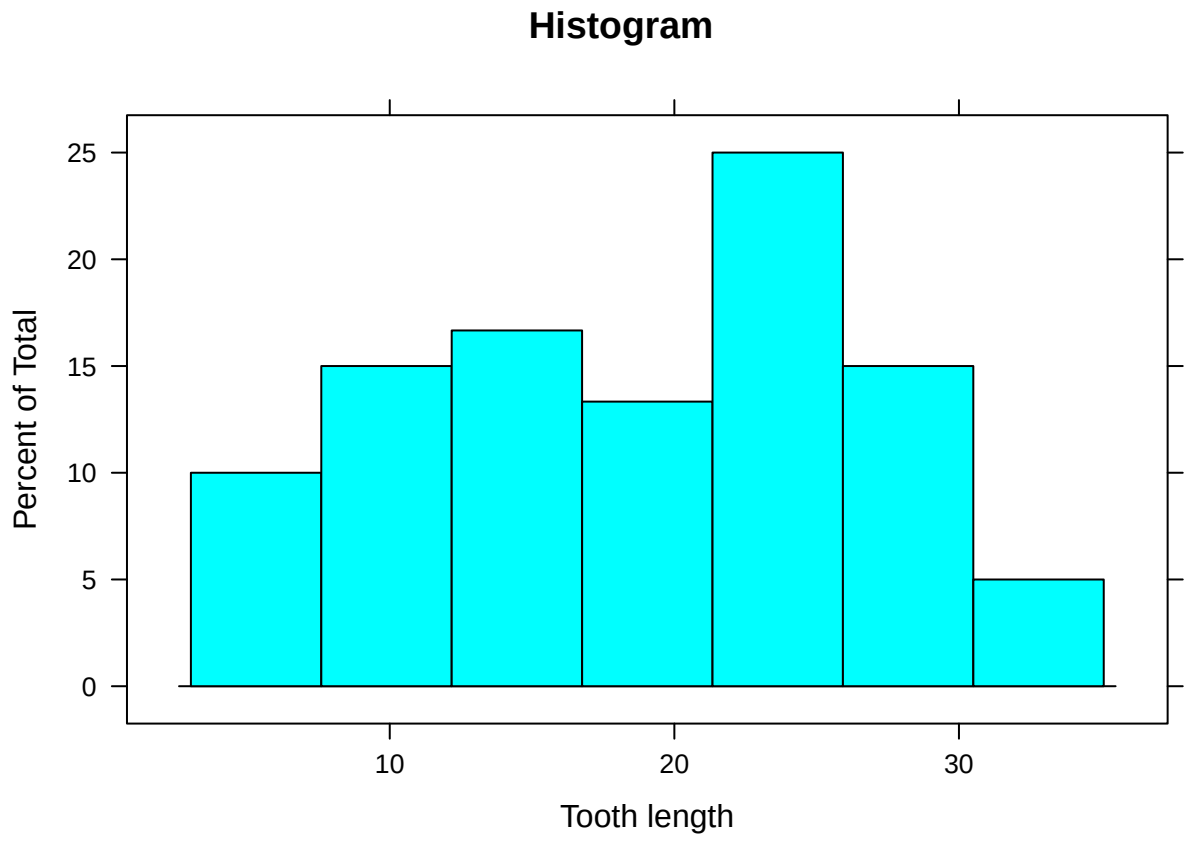
```
# mean by dose and supplement
mag <- aggregate(len ~ ., data = data.fr, mean)
xtabs(len ~ ., data = mag)
```

```
##     dose
## supp   0.5     1     2
##   OJ 13.23 22.70 26.06
##   VC  7.98 16.77 26.14
```

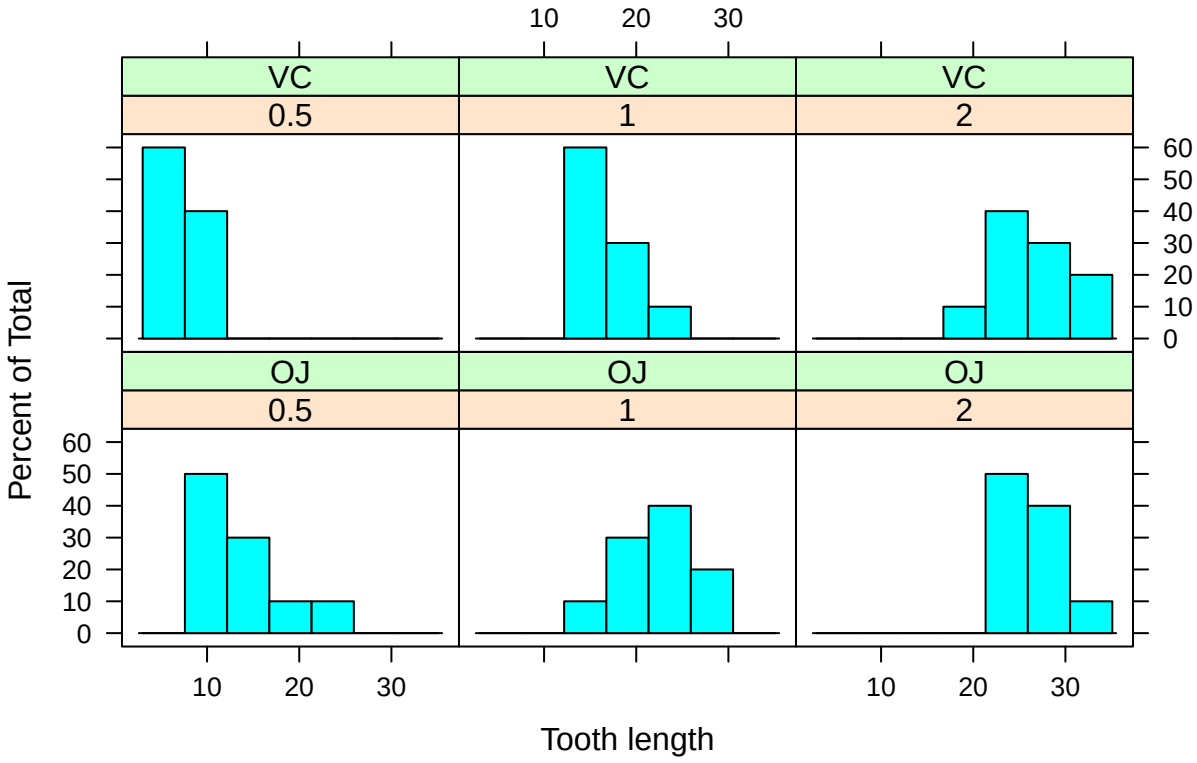
```r
# var by dose and supplement
vag <- aggregate(len ~ ., data = data.fr, var)
xtabs(len ~ ., data = vag)
```

```
##     dose
## supp       0.5         1         2
##   OJ 19.889000 15.295556  7.049333
##   VC  7.544000  6.326778 23.018222
```

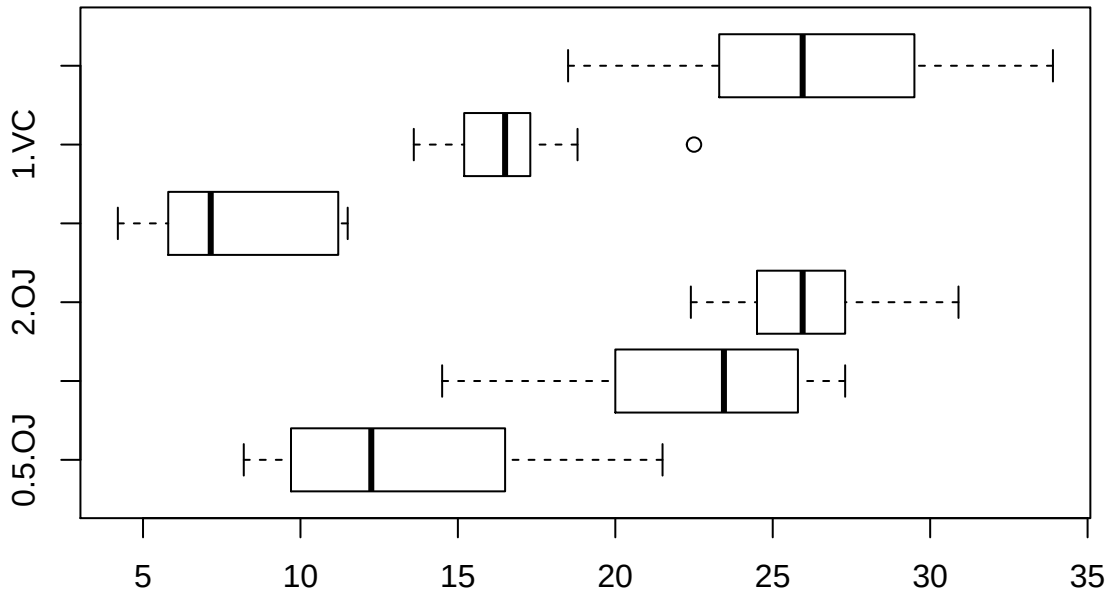
```r
# histogram of tooth growth
histogram( ~ data.fr$len, data = data.fr,
           main = "Histogram", xlab = "Tooth length")
```




```r
# histogram per dose and type of supplement
data.fr$dose <- as.factor(data.fr$dose)
histogram( ~ data.fr$len | data.fr$dose+data.fr$supp, data = data.fr,
           main = "Histogram of tooth length by Supplement and Dose",
           xlab = "Tooth length")
```




```
# and boxplot per dose and type of supplement
boxplot(data.fr$len ~ data.fr$dose + data.fr$supp, horizontal = TRUE)
```



The basic exploratory analysis suggests there are differences in the tooth length depending on the dose and type of supplement.

Confidence Intervals

We test the difference in the means of tooth length between the group that received one type of supplement and the other group.

```
# Independent groups T-test with unequal variance
# for assessing differences from supplement
t.test(data.fr$len~data.fr$supp, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  data.fr$len by data.fr$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
```

```r
# For assessing differences from dose, an ANOVA model would be better,
# however we are instructed to use T-tests:

# 0.5 vs rest
data.fr$d05 <- 0
data.fr$d05[which(data.fr$dose == "0.5")] <- 1
aggregate(len ~ d05, data = data.fr, var)
```

```
##   d05      len
## 1   0 26.82558
## 2   1 20.24787
```

```r
t.test(data.fr$len~data.fr$d05, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  data.fr$len by data.fr$d05
## t = 9.4907, df = 43.261, p-value = 3.873e-12
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   9.696667 14.928333
## sample estimates:
## mean in group 0 mean in group 1
##         22.9175         10.6050
```

```r
# 1.0 vs rest
data.fr$d10 <- 0
data.fr$d10[which(data.fr$dose == "1")] <- 1
aggregate(len ~ d10, data = data.fr, var)
```

```
##   d10      len
## 1   0 78.36666
## 2   1 19.49608
```

```r
t.test(data.fr$len~data.fr$d10, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  data.fr$len by data.fr$d10
## t = -0.8071, df = 57.994, p-value = 0.4229
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.811216  2.046216
## sample estimates:
## mean in group 0 mean in group 1
##         18.3525         19.7350
```

```r
# 2.0 vs rest
data.fr$d20 <- 0
data.fr$d20[which(data.fr$dose == "2")] <- 1
aggregate(len ~ d20, data = data.fr, var)
```

```
##   d20      len
## 1   0 40.73600
## 2   1 14.24421
```

```r
t.test(data.fr$len~data.fr$d20, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  data.fr$len by data.fr$d20
## t = -8.3085, df = 56.202, p-value = 2.347e-11
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.565108  -8.294892
## sample estimates:
## mean in group 0 mean in group 1
##           15.17           26.10
```

Conclusions

- With a 95% confidence, there is no conclusive evidence the type of supplement explains the difference in tooth growth.
- From the p-value, when considering the type of supplement, differences in tooth growth could be considered at significance levels above 6.1%.
- From the p-value, when considering the dose, differences in tooth growth could be considered from 0.5mg vs the rest of doses.
- There is no significant difference in tooth growth when considering a 1mg dose vs the rest of doses.
- From the p-value, when considering the dose, differences in tooth growth could be considered from 2.0mg vs the rest of doses.