Analysis of ToothGrowth Data

E Drake

Exploratory Data Analyses

We start by loading the data set and viewing it's structure using the str function:

```
data(ToothGrowth)
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 ...
```

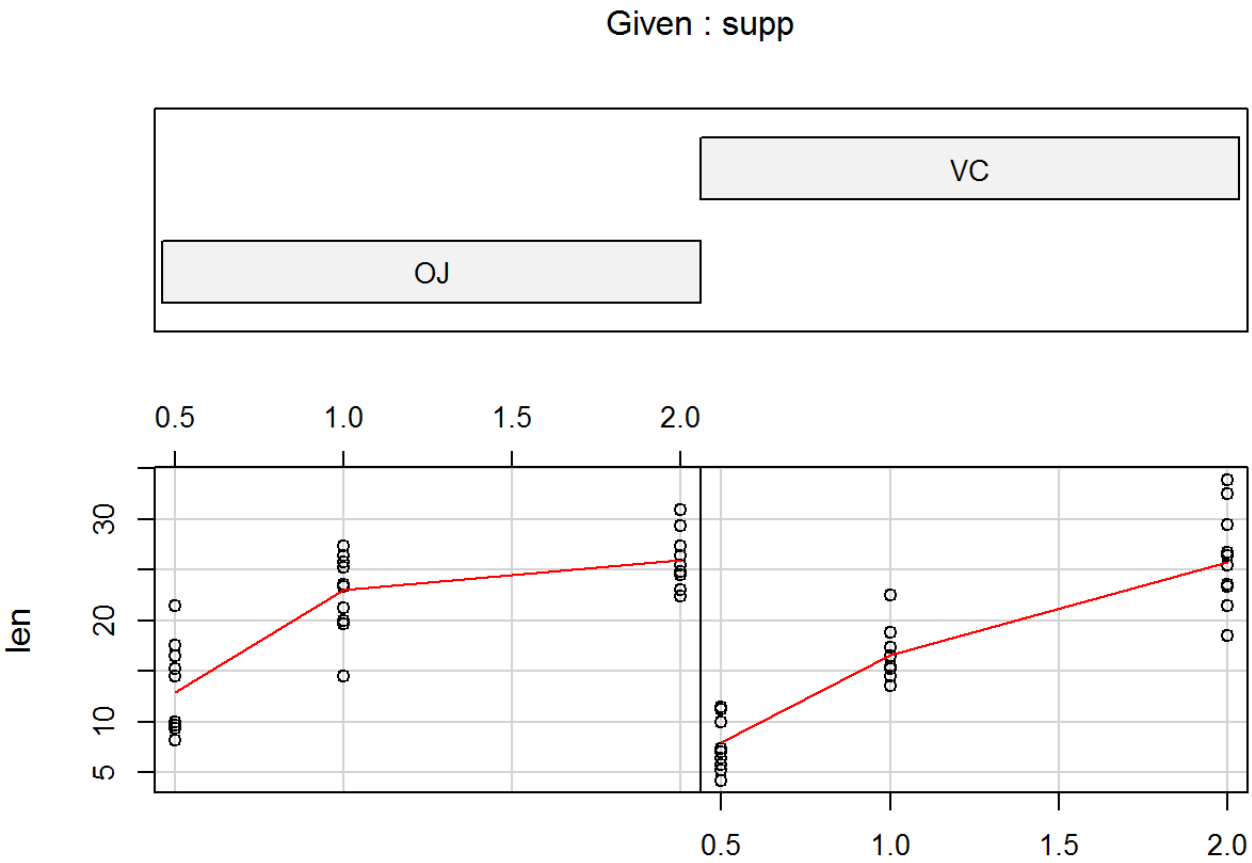
The output is telling us that this is a data frame with three variables.

Cross referencing with the documentation, we understand that the data shows the results of administering two different vitamin C supplements (orange juice and ascorbic acid) each at three different dose levels (0.5, 1 and 2 mg).

```
?ToothGrowth
```

We can plot an overview of the data using the following code (as suggested in the R documentation):

```
require(graphics)
coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth,
       xlab = "ToothGrowth data: length vs dose, given type of supplement")
```



ToothGrowth data: length vs dose, given type of supplement

To summarise the results of the intial analysis:

- There are 10 data points for each of 6 supplement - dose combination.
- These are 6 independent groups - the experiment would not make sense if the same 10 guinea pigs had been exposed to all 6 supplement - dose combinations.
- The graph seems to show that higher doses mean greater tooth length and that orange juice is more effective than ascorbic acid. In the next section we will test then statistical significance of these trends.

Hypothesis Tests

We are going to treat each of the six supplement - dose combinations separately. I believe this will give more insight than grouping the data across supplement or across dose.

I'll first create 6 lists that store the tooth length for a single supplement - dose combination:

```
vc05 <- ToothGrowth$len[1:10] # Ascorbic acid 0.5 mg
vc10 <- ToothGrowth$len[11:20] # Ascorbic acid 1.0 mg
vc20 <- ToothGrowth$len[21:30] # Ascorbic acid 2.0 mg
oj05 <- ToothGrowth$len[31:40] # Orange juice 0.5 mg
oj10 <- ToothGrowth$len[41:50] # Orange juice 1.0 mg
oj20 <- ToothGrowth$len[51:60] # Orange juice 2.0 mg
```

Now let's see if there is any differences between the groups:

Differences between doses of the same supplement:

H0: There is no difference between tooth length between ascorbic acid dosed at 0.5mg and at 1.0mg:

```
t.test(vc05,vc10, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  vc05 and vc10
## t = -7.4634, df = 18, p-value = 6.492e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.264346  -6.315654
## sample estimates:
## mean of x mean of y
##      7.98     16.77
```

We see in the R output that the p-value is very small. With the confidence interval being [-11.3, -6.3] we can reject the null hypothesis and instead conclude:

An ascorbic acid dose of 1.0mg is linked to a significantly greater tooth length than an ascorbic acid dose of 0.5mg.

H0: There is no difference between tooth length between ascorbic acid dosed at 1.0mg and at 2.0mg:

```
t.test(vc10,vc20, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  vc10 and vc20
## t = -5.4698, df = 18, p-value = 3.398e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -12.96896  -5.77104
## sample estimates:
## mean of x mean of y
##     16.77     26.14
```

Once again, the p-value is very small. We reject the null hypothesis and conclude instead:

An ascorbic acid dose of 2.0mg is linked to a significantly greater tooth length than an ascorbic acid dose of 1.0mg.

Without further statistical analysis we can now also conclude that:

An ascorbic acid dose of 2.0mg is linked to a significantly greater tooth length than an ascorbic acid dose of 0.5mg.

We can conduct similar analysis on the orange juice doses:

```
t.test(oj05,oj10, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  oj05 and oj10
## t = -5.0486, df = 18, p-value = 8.358e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.410814  -5.529186
## sample estimates:
## mean of x mean of y
##     13.23     22.70
```

```
t.test(oj10,oj20, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  oj10 and oj20
## t = -2.2478, df = 18, p-value = 0.03736
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -6.5005017 -0.2194983
## sample estimates:
## mean of x mean of y
##     22.70     26.06
```

Our conclusions are:

An orange juice dose of 1.0mg is linked to a significantly greater tooth length than an orange dose of 0.5mg.

An orange juice dose of 2.0mg is linked to a significantly greater tooth length than an orange dose of 1.0mg.

An orange juice dose of 2.0mg is linked to a significantly greater tooth length than an orange dose of 0.5mg.

Differences between supplements at the same dose:

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

```
##
##  Two Sample t-test
##
## data:  oj05 and vc05
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.770262 8.729738
## sample estimates:
## mean of x mean of y
##     13.23      7.98
```

```
t.test(oj10,vc10, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  oj10 and vc10
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.840692 9.019308
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

```
t.test(oj20,vc20, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  oj20 and vc20
## t = -0.0461, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

These tests tell us:

At doses of 0.5mg and 1.0mg tooth length is greater when the supplement is orange juice rather than ascorbic acid.

There is no evidence to suggest a difference in tooth length for orange juice at 2.0mg compared to ascorbic acid at 2.0mg.

Conclusions:

- **For both ascorbic acid and orange juice supplements, a higher dose of Vitamin C is linked to longer teeth in guinea pigs.**
- **At doses of 0.5mg and 1.0mg the supplement orange juice is linked to longer teeth than the supplement of ascorbic acid.**
- **At a dose of 2.0mg, there is no significant difference in tooth length between the two supplements.**

Assumptions:

- The 6 groups of guinea pigs were independent, randomly assigned and kept in identical conditions except for supplement and dose of Vitamin C.
- The sample of 60 guinea pigs are representative of the entire population of guinea pigs.
- The variance of tooth length is equal between each of the 6 groups.