Statistical Inference: Tooth Growth Analysis

Smeths

19 August 2015

Summary

The aim of this project is to

1) Perform an exploratory analysis of the data set, so as to understand it.
2) Clearly state assumptions made about the data and hence produce confidence intervals/hypothesis tests inline with the assumptions made
3) Form a conclusion based on this analysis

loading data

```
# Loading the data and getting some initial information
data(ToothGrowth)
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
```

```
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
```

Exploratory Analysis

```
library(ggplot2)
# Producing plots of len against dose with a factor of supp
g <- ggplot(ToothGrowth,aes(dose,len))
g <- g + geom_point(color = "steelblue")
g <- g + labs(x="dose",
              y="length")
g <- g + facet_grid(.~supp)
g
```
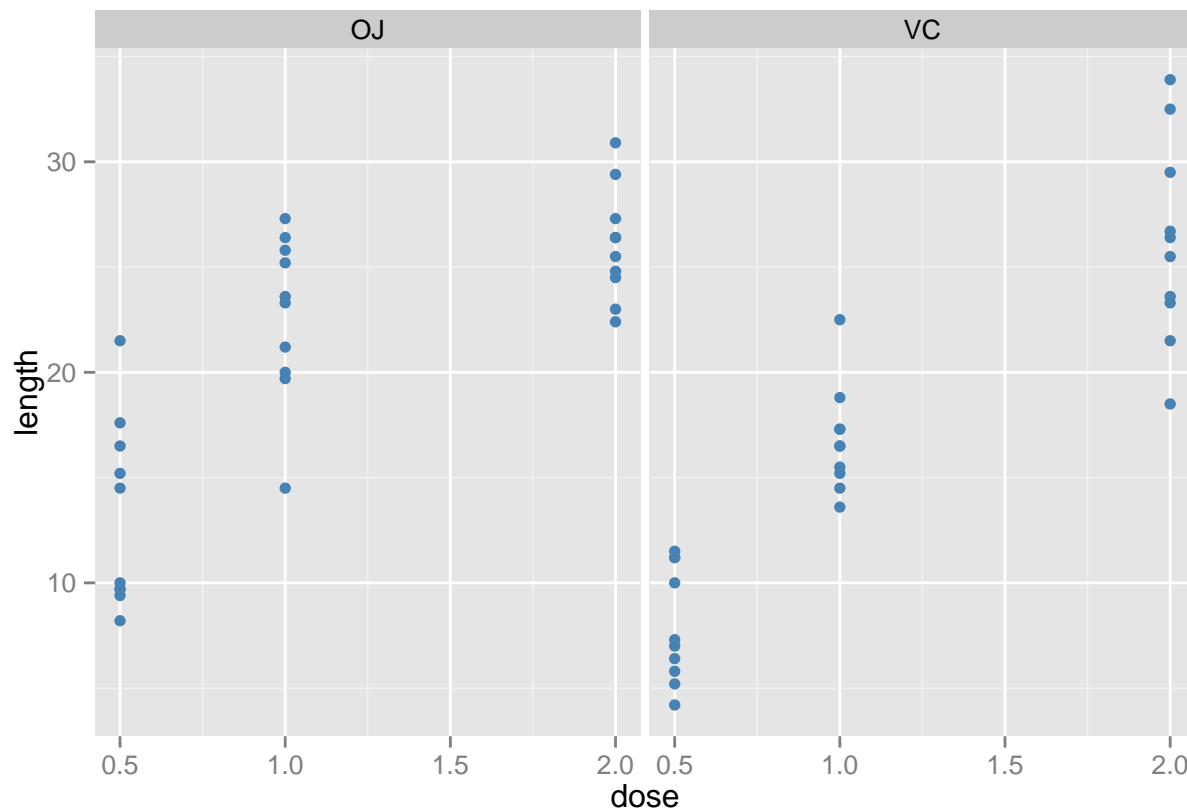
The plots clearly show that increasing the dose of OJ or VC increases the len (preassumably tooth length variable). On the whole it seems that the OJ data is above the VC data, so it would be worth testing at the 95% level for the different doses to see if this is true.

Confidence Intervals

The following assumptions shell be made about the data

1) The OJ and VC data is not paired
2) The sample size (10 for each dose/supplement) is small so a t distribution will be used
3) The variance is not constant

```
# performing t test for dose = 0.5
t.test(len~supp,
       paired = FALSE,
       var.equal = FALSE,
       data = ToothGrowth[ToothGrowth$dose == 0.5,])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
```

2

```
## sample estimates:
## mean in group OJ mean in group VC
##           13.23              7.98
```

```
# performing t test for dose = 1.0
t.test(len~supp,
       paired = FALSE,
       var.equal = FALSE,
       data = ToothGrowth[ToothGrowth$dose == 1.0,])
```

```
##
##   Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##           22.70             16.77
```

```
# performing t test for dose = 2.0
t.test(len~supp,
       paired = FALSE,
       var.equal = FALSE,
       data = ToothGrowth[ToothGrowth$dose == 2.0,])
```

```
##
##   Welch Two Sample t-test
##
## data:  len by supp
## t = -0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##           26.06             26.14
```

Conclusion

The confidence intervals generated do not contain 0 for doses of 0.5 and 1.0, which means we can conclude there is significant evidence that supplement OJ causes a greater increase in tooth length for these doses. However, for a dose of 2.0 the confidence interval does contain 0 and therefore there is not significant evidence to suggest that supplement OJ or VC causes a greater increase in tooth length for this dose.