Tooth Growth

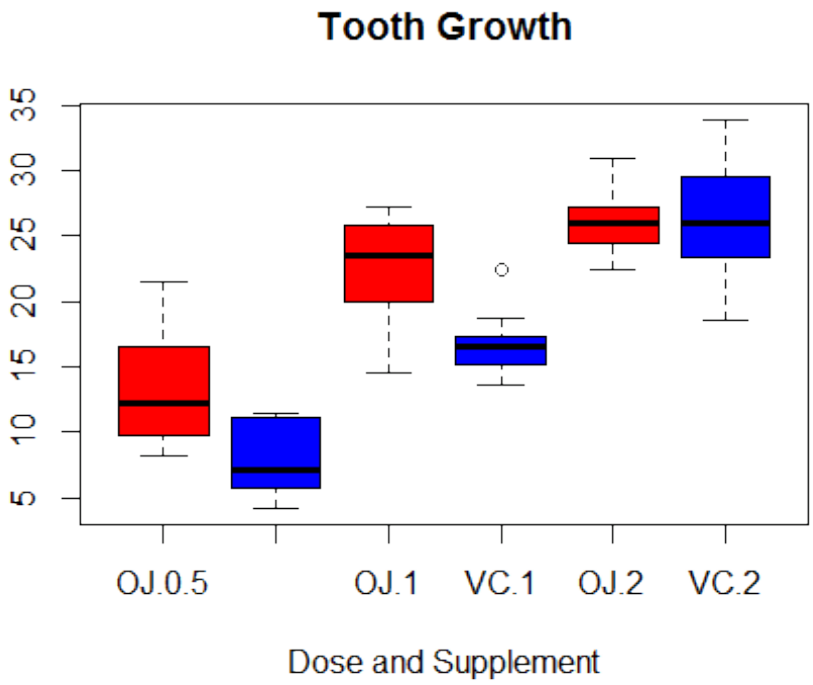
Candice MH

Sunday, November 09, 2014

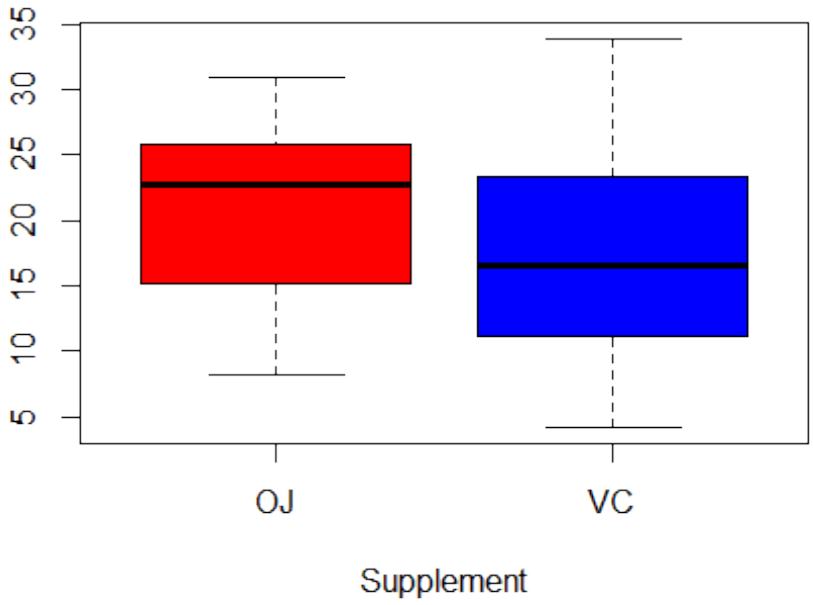
This project explores the ToothGrowth data in the R datasets package, by performing some basic exploratory data analyses, summarising the data, and using confidence intervals and hypothesis tests to compare tooth growth by supp and dose. Conclusions and assumptions will be stated throughout.

A summary of the tooth growth data is performed. The response variable Tooth Growth (len) varies between 4 and 34, and there are two explanatory variables: dose (taking values 0.5, 1 and 2) and supplement (taking values OJ and VC).

An exploration of the relationship between tooth growth, and dose and supplements, is achieved through side-by-side box plots of tooth growth for each combination of dose and supplement, as shown below. In general, higher doses seem to produce greater tooth growth, and OJ seems to produce greater tooth growth (except when dose equals 2). The plot reveals heterogeneity.



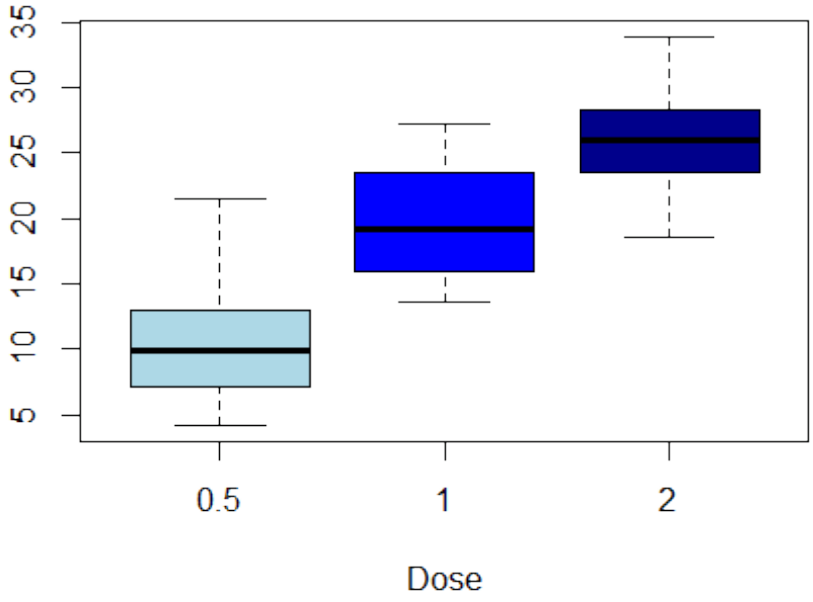
A visual exploration of the impact of supplement on tooth growth is given in the box plot below. Supplement OJ appears to have a higher mean tooth growth than supplement VC.



A one-sided two-sample t-test is performed to test for a significantly greater mean (H0: tooth growth (len) is equal for both supplements, versus H1: tooth growth (len) is greater for OJ than VC). The assumptions of independence, equal variances (ratio of variances is 1.57) and normality appear to be met, and the resultant p-value is 0.03032. We have strong evidence to reject the null hypothesis and conclude that tooth growth has a greater length given supplement OJ compared to VC.

```
##
##  Welch Two Sample t-test
##
## data:  len[supp == "OJ"] and len[supp == "VC"]
## t = 1.915, df = 55.31, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4683    Inf
## sample estimates:
## mean of x mean of y
##     20.66     16.96
```

A visual exploration of the impact of dose on tooth growth is given in the box plot below. Higher doses seem to lead to greater mean tooth growth.



Given that ANOVA is not covered in this course, I've tested for significantly greater means for greater doses through two t-tests, after adjusting for a family wise error rate of 0.05 using a Bonferroni correction. The assumptions underpinning two-sample t-tests appear to be met as the variances are approximately equal (ratio of largest to smallest variance is 1.42), the data is roughly normal (as shown in the above barplots), and the samples are independent.

The first one-sided two-sample t-test is that H1: dose=1 results in a higher mean tooth growth than dose=0.5 (against H0: mean tooth growth is equal for dose=1 and dose=0.5). The resulting p-value 6.342e-08 means we have very strong evidence to reject the null hypothesis and conclude that dose=1 results in greater mean tooth growth than dose=0.5.

The second one-sided two-sample t-test is that H1: dose=2 results in a higher mean tooth growth than dose=1 (against H0: mean tooth growth is equal for dose=2 and dose=1). The resulting p-value 9.532e-06 means that we have very strong evidence to reject the null hypothesis and conclude that dose=2 results in greater mean tooth growth than dose=1.

R output for these t-tests are provided in the appendix.