Analyze of the ToothGrowth data

Pratheep

Overview

The tooth growth dataset contains the effect of vitamin C in Tooth growth in Guinea pigs.The response is the length of odontoblasts (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).

Synopsis

The load Growth data will be loaded the some some basic exploratory data analyses will be performed including basic summary of the data. Thereafter confidence intervals and hypothesis tests will be used to compare tooth growth by supp and dose. Finally conclusions will be stated together with the assumptions needed.

Load

```
data(ToothGrowth)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

Basic analysis

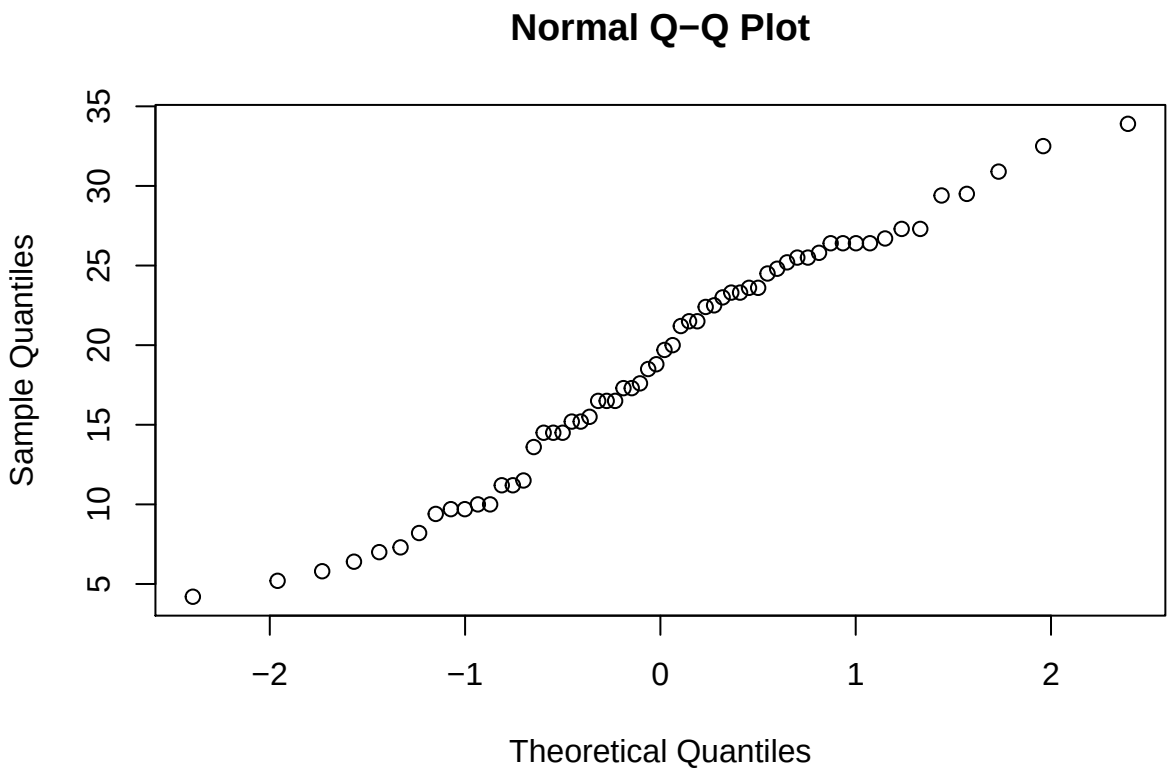
```
summary(ToothGrowth)
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

```
#Check for the nmber of doages.
table(ToothGrowth$dose)
```

```
##
## 0.5   1   2
##  20  20  20
```

```
#Check for normality of 'len' dstribution.
qqnorm(ToothGrowth$len)
```

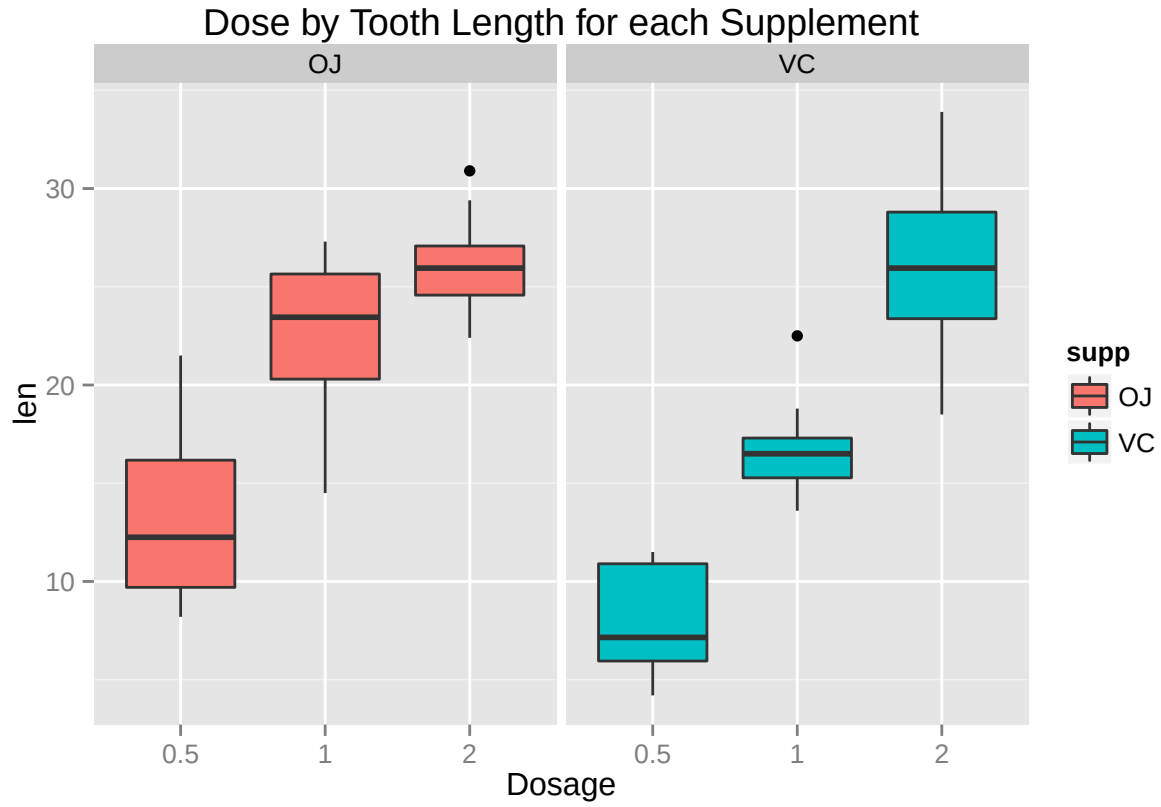


length is mostly distributed normally, thus methods for normal distribution will be used for further anlysis. Same number of instance for all dosage types, which allows balanced analysis. Being mutivariate, Anova

```
#Compare effects
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.2
```

```
ggplot(ToothGrowth, aes(x=dose, y=len,fill=supp))+
    geom_boxplot()+ facet_grid(.~supp)+ labs(x="Dosage")+
    ggtitle("Dose by Tooth Length for each Supplement")
```



An Anova test confirms the differences in improvements.

```
aov.result <- aov(len ~ supp*dose, data=ToothGrowth)
summary(aov.result)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## supp         1  205.4   205.4  15.572 0.000231 ***
## dose         2 2426.4  1213.2  92.000  < 2e-16 ***
## supp:dose    2  108.3    54.2   4.107 0.021860 *
## Residuals   54  712.1    13.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As seen both sup and does are significant at 0.0001 level, rejecting the null hypothesis. The interaction of supp and dose is significant with 95% confidence. (0.05 level)

The confidence interval for the means.

```
confint(aov.result)
```

```
##                   2.5 %    97.5 %
## (Intercept)  10.9276907 15.532309
## suppVC       -8.5059571 -1.994043
## dose1         6.2140429 12.725957
## dose2         9.5740429 16.085957
## suppVC:dose1 -5.2846186  3.924619
## suppVC:dose2  0.7253814  9.934619
```

Conclusion

Both supplements seem to have positive effect on tooth growth. Furthermore, at higher doses the improvements are muchbetter, with orange juice having less variance than vc.

The analysis is based on two assumptions: 1) The doseage and the suppliment were randomly applied on the Guinea pigs 2) The mean is normally distributed.