# Statistical Inference Project2

Yujing Cao

Sunday, September 21, 2014

PART 1 Load Dataset "ToothGrowth"

```
dataset=ToothGrowth
library(ggplot2)
```

PART 2 Basic Exploratory Data Analysis

```
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
  geom_bar(stat="identity",) +
  facet_grid(. ~ supp) +
  xlab("Dose in miligrams") +
  ylab("Tooth length") +
  guides(fill=guide_legend(title="Supplement type"))
```
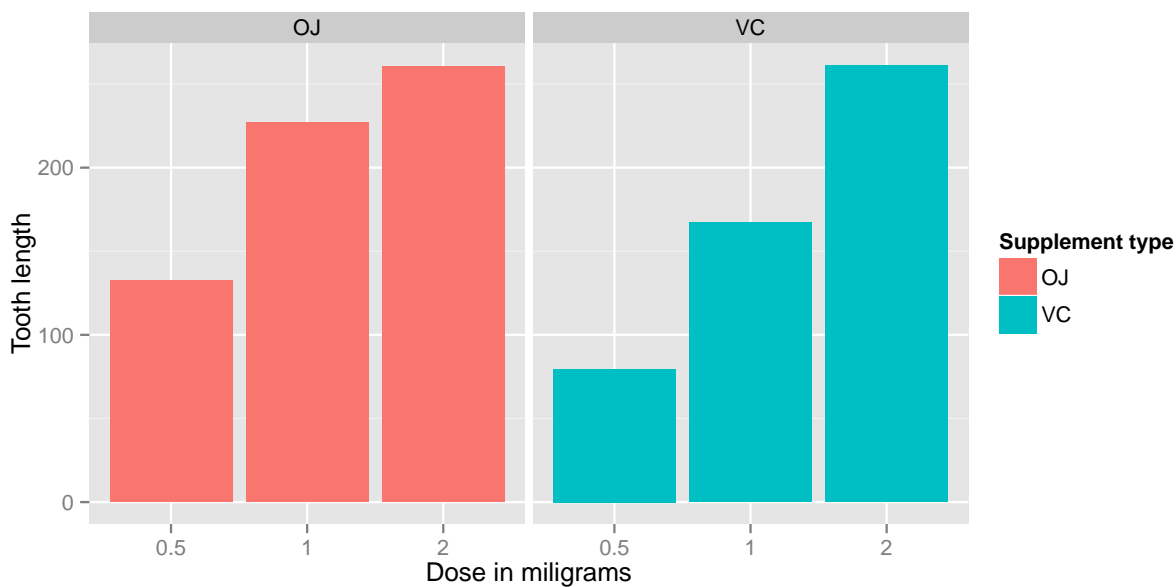

As we can from the graph, there is a positive correlation between the tooth length and the dose of Vitamin C for both delivery methods.In the other word, as the dose of Vitamin C increases, the tooth length increases.

```
fit <- lm(len ~ dose + supp, data=ToothGrowth)
summary(fit)
```

```
##
## Call:
## lm(formula = len ~ dose + supp, data = ToothGrowth)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.600 -3.700  0.373  2.116  8.800
```

```
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.273      1.282    7.23  1.3e-09 ***
## dose           9.764      0.877   11.14  6.3e-16 ***
## suppVC        -3.700      1.094   -3.38   0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 4.24 on 57 degrees of freedom
## Multiple R-squared:  0.704,  Adjusted R-squared:  0.693
## F-statistic: 67.7 on 2 and 57 DF,  p-value: 8.72e-16
```

This summary shows us that the adjust R-squared is 0.6934 which means the model explains 70% of the variance in the data. The interceptis 9.2725,meaning without supplyment Vitamin C, the average tooth growth is 9.2725 unit. The coefficient of dose is 9.7626, which means the tooth growth increases 9.7636 units with increasing the dose 1 miligrams.Since suppVC is categorical data, dummy variables are used. When ascorbic acid are given, the tooth growth will decrease 5.889905 units; if the orange juice are given, the tooth growth will increase 5.889905 units.

PART 3 Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose

```r
confint(fit)
```

```
##             2.5 % 97.5 %
## (Intercept)  6.705  11.84
## dose         8.008  11.52
## suppVC      -5.890  -1.51
```

The 95% of confidence interval of dose is [8.007741,11.519402], and 95% of confidence interval of suppVC is [-5.889905,-1.510095]. The null hypothesis is coefficients are 0, that is, dose and suppVC don't have significant effect on tooth growth. However, neither of these two coefficient cover 0, thus, null hypothesis is rejected given the significance level is 5%.