Statistical Inference Project

Mayank Pundhir

October 26, 2014

Part 2 of the Project: Basic inferential data analysis

In this Basic Inferential data analysis, ToothGrowth dataset from R has been investigated which contains data of length of teeth of guinea pigs with different levels of Vitamin C. All questions were answered.

1. Load the ToothGrowth data and perform some basic exploratory data analyses

Ans:

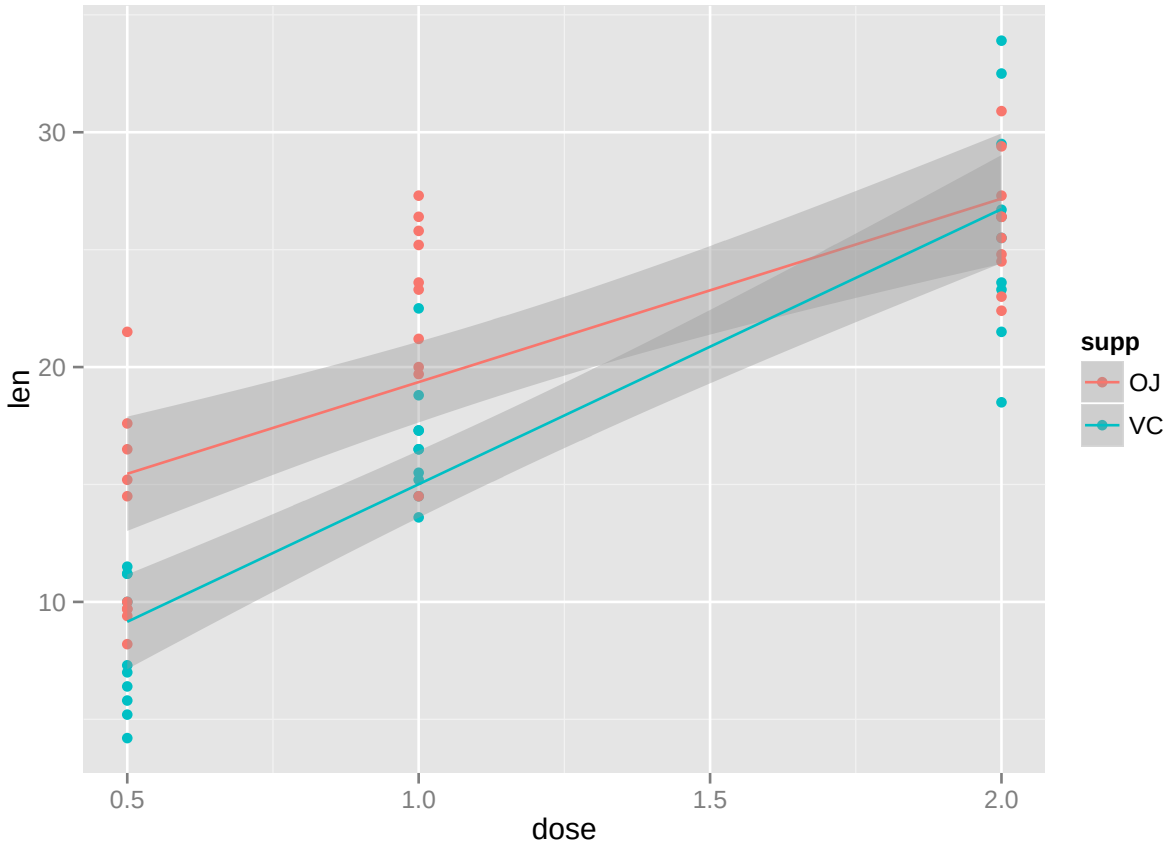
```
data(ToothGrowth)
str(ToothGrowth)

## '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 ...
```

A graph was plotted which showed the increase of length of teeth with increase in dosage of Vitamin C. Graph also suggested more length from Oranje Juice than Ascorbic Acid.

2. Provide a basic summary of the data.

Ans: It contains 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).

```
library(ggplot2)
ggplot(ToothGrowth) + geom_point(aes(x=dose, y=len, colour=supp))+
  geom_smooth(aes(x=dose, y=len,colour=supp), method="lm")
```



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

Ans: Several hypothesis were tested to measure the significance of dosage and supplemet type on length of teeth. Increase in length by increasing Vitamin C dosage from 0.5 mg to 1 mg was 9.13 whereas further increase of dosage to 2 mg only results in increase in 6.365. The difference from different supplement method was 3.7. Confidence Intervals of the estimates are given in Appendix.

4. State your conclusions and the assumptions needed for your conclusions.

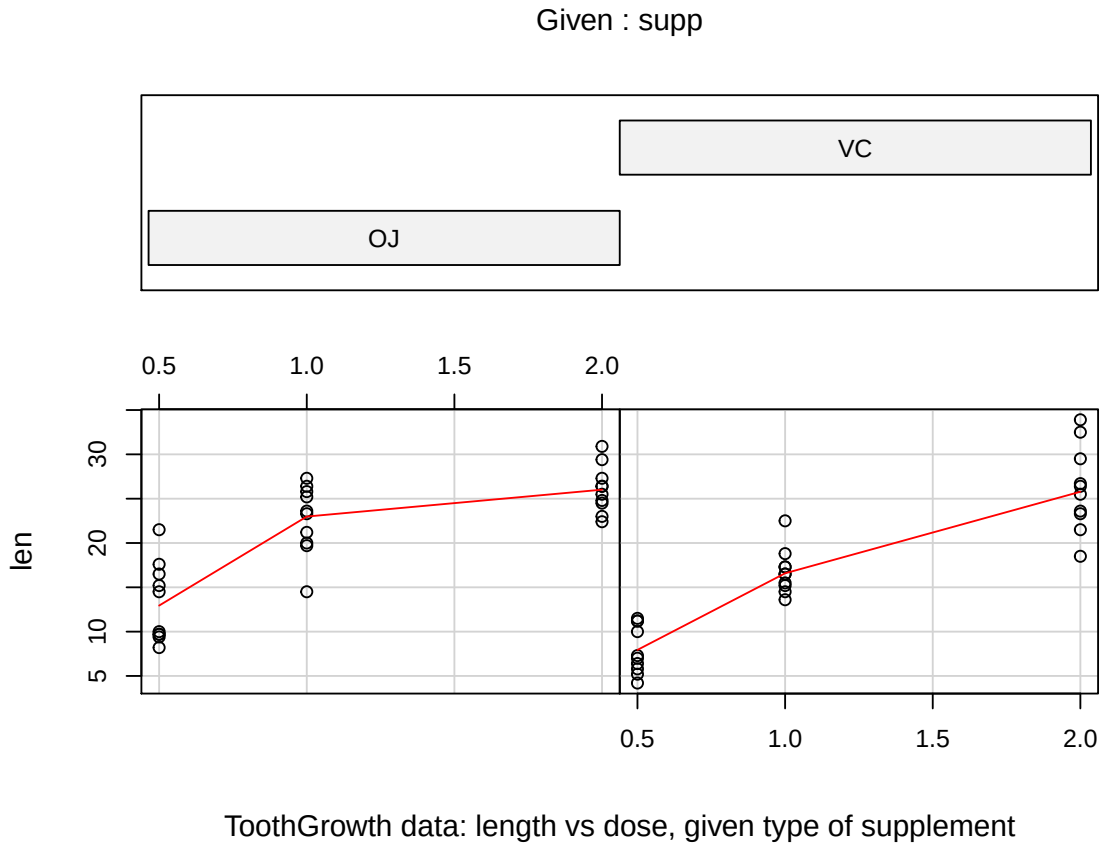
Ans: The tests showed there is a strong relationship between dosage of Vitamin C and length of teeth. They also suggested the significant increase in length when the delivery method was Oranje Juice. Interactions between dosage and delivery methods were assumed to be negligible and therefore not taken into account. Anova shows that Residuals are significant. The interaction between dosage and supp was also not much significant.

Appendix

Part 2:

Graph of the data.

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



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

```
a1<-ToothGrowth[ToothGrowth$dose==0.5,1];
a2<-ToothGrowth[ToothGrowth$dose==1,1]
a3<-ToothGrowth[ToothGrowth$dose==2,1];
t.test(a2-a1)

##
##  One Sample t-test
##
## data:  a2 - a1
## t = 6.967, df = 19, p-value = 1.225e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   6.387 11.873
## sample estimates:
## mean of x
##      9.13
```

```
t.test(a3-a2)

##
##  One Sample t-test
##
## data:  a3 - a2
## t = 4.605, df = 19, p-value = 0.0001934
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.472 9.258
## sample estimates:
## mean of x
##     6.365
```

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

```
b1<-ToothGrowth[ToothGrowth$supp=="VC",1];
b2<-ToothGrowth[ToothGrowth$supp=="OJ",1]
t.test(b2-b1)

##
##  One Sample t-test
##
## data:  b2 - b1
## t = 3.303, df = 29, p-value = 0.00255
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.409 5.991
## sample estimates:
## mean of x
##       3.7
```

ANOVA Table was also created

```
aov(len~dose*supp,ToothGrowth)

## Call:
##    aov(formula = len ~ dose * supp, data = ToothGrowth)
##
## Terms:
##                    dose   supp dose:supp Residuals
## Sum of Squares  2224.3  205.3      88.9     933.6
## Deg. of Freedom      1      1         1        56
##
## Residual standard error: 4.083
## Estimated effects may be unbalanced
```