Tooth Growth Data Analysis

Patricio Del Boca

27/09/2015

Overview

- Load the ToothGrowth data and perform some basic exploratory data analyses
- Provide a basic summary of the data.
- Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
- State your conclusions and the assumptions needed for your conclusions.

Load Data

```
library(ggplot2)
library(datasets)
require(graphics)
data(ToothGrowth)
toothGrowth <- ToothGrowth
toothGrowth$dose <- as.factor(toothGrowth$dose) # convert to factor
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
```

Exploratory Data Analysis

Number of observations for each treatment/experiment. Total guinea pigs = 60:

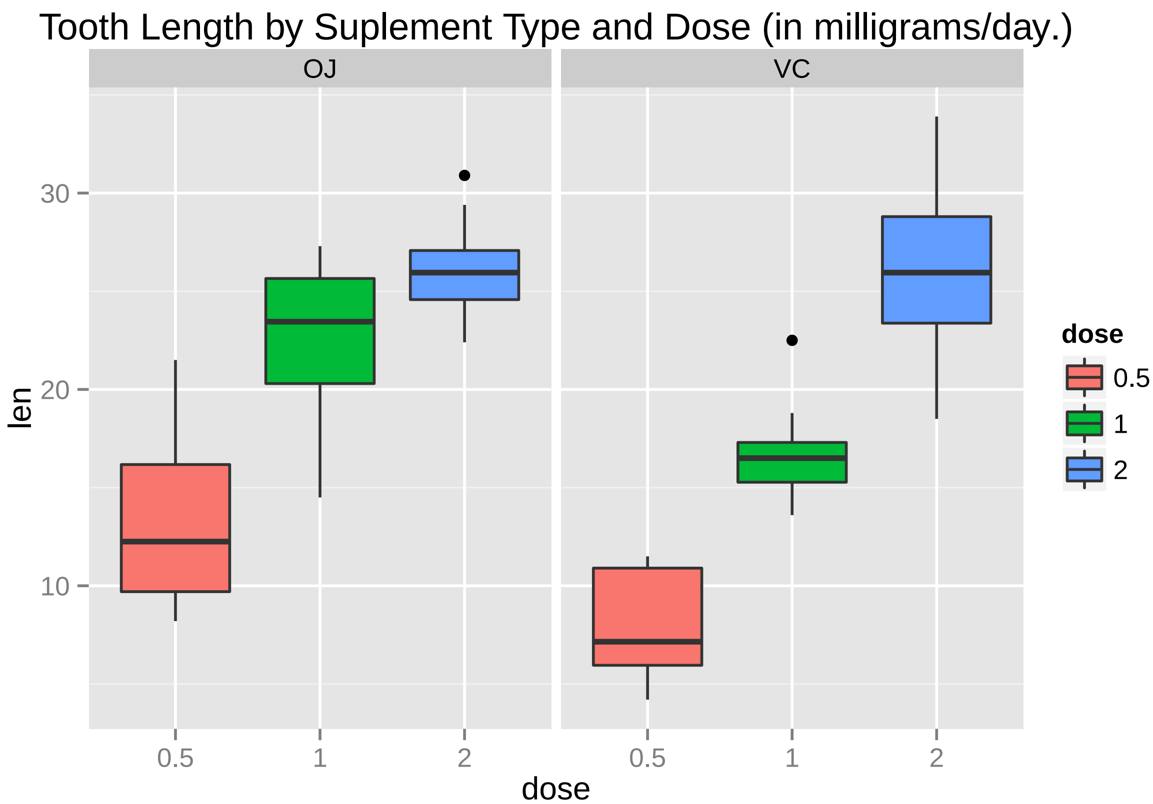
```
table(toothGrowth$supp, toothGrowth$dose)
```

```
##
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

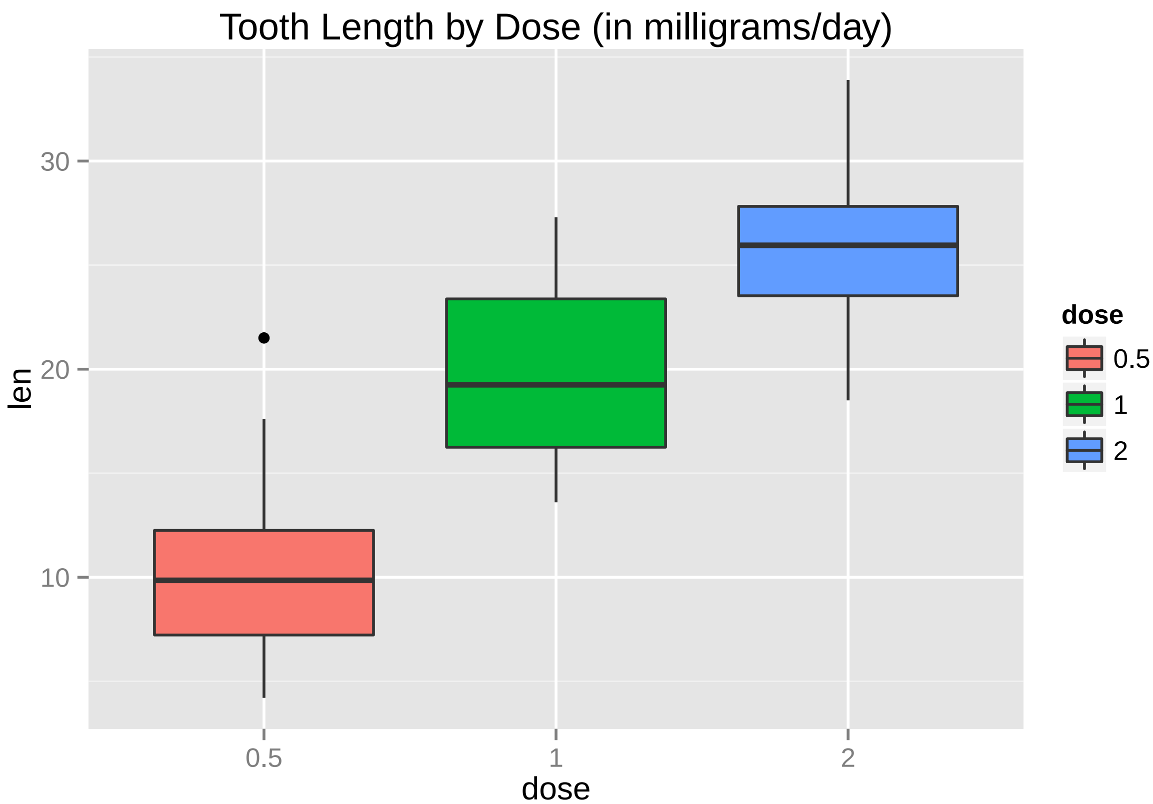
Some plots to explore data:

```
ggplot(data=toothGrowth, aes(x=supp,y=len,fill=supp)) +
  geom_boxplot() +
  ggtitle("Tooth Length by Supplement applied.")
```

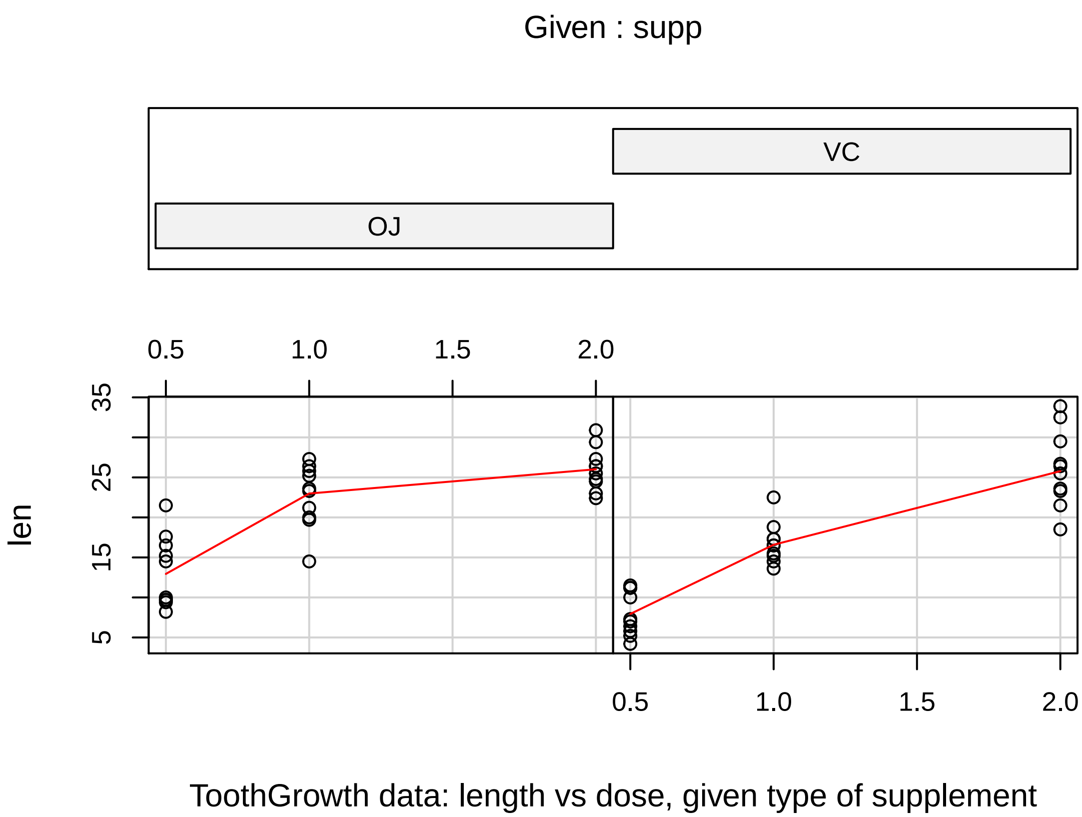

```
ggplot(data=toothGrowth, aes(x=dose,y=len,fill=dose)) +
  geom_boxplot() +
  facet_grid(.~supp) +
    ggtitle("Tooth Length by Suplement Type and Dose (in milligrams/day.)")
```




```
ggplot(data=toothGrowth, aes(x=dose,y=len,fill=dose)) +
  geom_boxplot() +
    ggtitle("Tooth Length by Dose (in milligrams/day)")
```




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



In the grafic we can see that the difference is substantial when dosages are 0.5 and 1.0 respectively. However, when dosage is 2.0 there isn't a difference. Let's perform a T Test to corroborate this hipotesis.

Tests:

I will perform T Test for each subset of data separated by dosages

Assumptions:

- Samples are both unpaired and unequal variances.
- There aren't other cofounding factors
- All the guinea pigs have the same diet, size, care, etc

```
d0.5 <- toothGrowth[toothGrowth$dose == 0.5, ]
d1.0 <- toothGrowth[toothGrowth$dose == 1.0, ]
d2.0 <- toothGrowth[toothGrowth$dose == 2.0, ]
```

```
test0.5 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d0.5)
p0.5 <- test0.5$p.value
ci0.5 <- test0.5$conf
```

```
## [1] "P Value for Dose = 0.5: 0.00635861"
```

```
## [1] "Confidence Interval for Dose = 0.5: 1.71906"
## [2] "Confidence Interval for Dose = 0.5: 8.78094"
```

```
test1.0 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d1.0)
p1.0 <- test1.0$p.value
ci1.0 <- test1.0$conf
```

```
## [1] "P Value for Dose = 1.0: 0.00103838"
```

```
## [1] "Confidence Interval for Dose = 1.0: 2.80215"
## [2] "Confidence Interval for Dose = 1.0: 9.05785"
```

```
test2.0 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d2.0)
p2.0 <- test2.0$p.value
ci2.0 <- test2.0$conf
```

```
## [1] "P Value for Dose = 2.0: 0.963852"
```

```
## [1] "Confidence Interval for Dose = 2.0: -3.79807"
## [2] "Confidence Interval for Dose = 2.0: 3.63807"
```

Conclusions

P-values are lower than 0.05 for dose 0.5 and 1.0 but greater than for 2.0. We can say the supplements have different effecs for dosages 0.5 and 1.0 but there isn't enought evidence to claim so when the dossage is 2.0.